EuroWordNet General Document[1]

Version 3
Final
July 19, 1999

Piek Vossen (ed.)
University of Amsterdam
e-mail: piek.vossen@hum.uva.nl
http://www.hum.uva.nl/~ewn

[1] This document is an extension and enhancement of various other documents that have been published. For an overview of all publications on EuroWordNet see http://www.hum.uva.nl/~ewn.

Table of Contents

1. Introduction ... 5
2. Design of the multilingual database ... 8
   2.1. The Database Modules ... 8
   2.2. The Language Internal Relations ... 11
      2.2.1. Criteria for the identification of relations between synsets ... 12
      2.2.2. Relation Labels ... 13
         2.2.2.1. Conjunction /Disjunction ... 14
         2.2.2.2. Factivity ... 14
         2.2.2.3. Reversed ... 14
         2.2.2.4. Negation ... 15
      2.2.3. The subtypes of language-internal relations ... 15
         2.2.3.1. Synonymy ... 18
         2.2.3.2. Hyponymy ... 21
         2.2.3.3. Antonymy ... 24
         2.2.3.4. Meronymy ... 26
         2.2.3.5. ROLE and INVOLVED ... 28
         2.2.3.6. CO_ROLE ... 31
         2.2.3.7. CAUSES and IS_CAUSED_BY ... 32
         2.2.3.8. HAS_SUBEVENT and IS_SUBEVENT_OF ... 35
         2.2.3.9. IN_MANNER and MANNER_OF ... 36
         2.2.3.10. BE_IN_STATE and STATE_OF ... 37
         2.2.3.11. Derivational relations ... 37
         2.2.3.12. Instance and Class ... 37
         2.2.3.13. Undefined Relations: “fuzzynyms” ... 37
   2.3. Multilinguality ... 38
      2.3.1 Equivalence relations ... 38
      2.3.2. Inter-Lingual-Index ... 39
         2.3.2.1. Extending the ILI with new concepts ... 39
         2.3.2.2. Creating a coarser level of differentiation in the ILI ... 40
      2.3.3. Accessing complex equivalence mappings ... 43
   2.4. Variant Information ... 46
   2.5. EuroWordNet Import/Export Format ... 47
      2.5.1. Import/Export format for synsets ... 47
      2.5.2. Import/Export format for ILI-records ... 49
      2.5.3. Import format for Top-Concepts and Domains ... 50
3. Methodology ... 52
   3.1. Expand/Merge approach ... 52
   3.2. Base Concepts ... 53
   3.3. Top Ontology ... 58
      3.3.1. Classification of 1st-Order-Entities ... 62
      3.3.2. The classification of 2ndOrderEntities ... 65
         3.3.2.1. SituationTypes ... 65
         3.3.2.2 SituationComponents ... 66
4. The EuroWordNet database ... 72
5. Description of the CD-Rom ... 76
Acknowledgements ... 78
References ... 79
Appendix I: Base Concepts Selected by four sites in EuroWordNet ... 82
Appendix II Top Ontology Classification of the Base Conceps ... 83
Appendix III: Top Concept Cluster Combinations for Base Concepts ... 98
Appendix IV EuroWordNet Optional Variant Information ... 106

List of Tables

Table 1: WordNet1.5 Relations 15
Table 2: Language Internal Relations between synsets in EuroWordNet 17
Table 3: Language-Internal Relations between other data types in EuroWordNet 18
Table 4: The Equivalence Relations in EuroWordNet 38
Table 5: Dutch, Spanish and Italian Synsets linked to senses of “office” in the ILI. 41
Table 6: Composite ILI-records 43
Table 7: Overview of mapping relations to the ILI 44
Table 8: Intersection-pairs of translations of English, Dutch, Spanish and Italian Base Concepts 54
Table 9: Complete Intersections of Base Concept Selections 54
Table 10: Overlap of EWN2 nouns and EWN1 nouns (905 CBCs) 55
Table 11: Overlap of EWN2 verbs and EWN1 verbs (239 CBCs) 55
Table 12: Selected and Rejected Base Concepts for each language 56
Table 13: Number of Common Base Concepts represented in the local wordnets 56
Table 14: Base Concept Gaps in at least two wordnets 56
Table 15: Total Set of classified Base Concepts 61
Table 16: Definitions for first order top concepts 64
Table 17: Mapping of WordNet1.5 Lexicographer's file codes to EuroWordNet top-concepts 71

List of Figures

Figure 1: Synsets related to “car” in its first sense in WordNet1.5 5
Figure 2. The global architecture of the EuroWordNet database 9
Figure 3: Parallel wordnet structures in EuroWordNet linked to the same ILI-records. 11
Figure 4: Spanish synset “oficina” with extended EQ_METONYM link to a Composite ILI-record for “office” 42
Figure 5: Many-to-many mappings of near synonyms of “apparatus” synsets to ILI-records. 44
Figure 6: Ways of “killing” lexicalized in Dutch and not in English 45
Figure 7: Morpho-syntactic variant features allowed in EuroWordNet 46
Figure 8: Usage labels for variants allowed in EuroWordNet 47
Figure 9: Global overview of steps in building EuroWordNet 53
Figure 10: General outline of the vocabularies 58
Figure 11: The EuroWordNet Top-Ontology 61
Figure 12: Lattice structure of the EuroWordNet top-ontology 63
Figure 13: Accessing separate wordnets and their equivalence links 73
Figure 14: Accessing different wordnets via the Inter-Lingual-Index 73
Figure 15: Accessing different wordnets via the Top-Ontology 74
Figure 16: Accessing different wordnets via the Domain hierarchy 74
Figure 17: Projecting Dutch “vehicles” (1 level) to the Spanish wordnet 75
Figure 18: Folder with the EuroWordNet database stores 77

List of Abbreviations

GB = English wordnet of additions to WordNet1.5	EWN = EuroWordNet
NL = Dutch wordnet	ILI = Inter-Lingual-Index
IT = Italian wordnet	ILIR = Inter-Lingual-Index-record
ES = Spanish wordnet	WM = word meaning
DE = German wordnet	BC = Base Concept
FR = French wordnet	TC = Top-Concept
EE = Estonian wordnet	LIR = Language-internal relations
CZ = Czech wordnet	POS = part-of-speech
WN15 = WordNet1.5	

List of contributors to EuroWordNet

Person	*Affiliation*	*Country*
Adriana Roventini	ILC-CNR Pisa	IT
Andreas Wagner	University of Tuebingen	DE
Anselmo Peñas	Universidad Nacional de Educación a Distancia	ES
Antonietta Alonge	ILC-CNR Pisa	IT
Antonio Sanchez Valderrabanos	Lernout & Hauspie	BE
Carol Peters	ILC-CNR Pisa	IT
Cintha Harjadi	University of Amsterdam	NL
Clara Soler	Universitat de Barcelona	ES
Claude de Loupy	Bertin Technologies	F
Claudia Kunze	University of Tuebingen	DE
Dominique Dutoit	Memodata	F
David Fernández	Universidad Nacional de Educación a Distancia	ES
Elisabetta Marinai	ILC-CNR Pisa	IT
Erhard Hinrich	University of Tuebingen	DE
Fernando López	Universidad Nacional de Educación a Distancia	ES
Francesca Bertagna	ILC-CNR Pisa	IT
Frédérique Segond	Xerox Research Centre	F
Geert Adriaens	Lernout & Hauspie	BE
Gerard Escudero Bakx	Universitat Politècnica de Catalunya	ES
German Rigau Claramunt	Universitat Politècnica de Catalunya	ES
Haldur Õim	University of Tartu	EE
Heili Orav	University of Tartu	EE
Helmut Feldweg	University of Tuebingen	DE
Horacio Rodríguez	Universitat Politècnica de Catalunya	ES
Ilse Cuypers	Lernout & Hauspie	BE
Irene Castellón Masalles	Universitat de Barcelona	ES
Irina Chugur	Universidad Nacional de Educación a Distancia	ES
Ivonne Peters	University of Sheffield	GB
Javier Farreres de la Morena	Universitat Politècnica de Catalunya	ES
Jordi Turmo	Universitat Politècnica de Catalunya	ES
Josep Carmona	Universitat Politècnica de Catalunya	ES
Julio Gonzalo	Universidad Nacional de Educación a Distancia	ES
Kadri Vider	University of Tartu	EE
Karel Pala	Masaryk University	CZ
Laia Palouzi	Universitat de Barcelona	ES
Laura Benítez	Universitat Politècnica de Catalunya	ES
Laura Bloksma	University of Amsterdam	NL
Laurent Catherin	LIA, Avignon	F
Laurent Griot	Xerox Research Centre	F
Leho Paldre	University of Tartu	EE
Linda Schippers	Lernout & Hauspie	BE
Lluis Màrquez	Universitat Politècnica de Catalunya	ES
Lluís Padró	Universitat Politècnica de Catalunya	ES
Luca Tarasi	ILC-CNR Pisa	IT
M. Antonia Martí	Universitat de Barcelona	ES
M.Felisa Verdejo	Universidad Nacional de Educación a Distancia	ES
Marc El-Beze	LIA, Avignon	F
Mariona Taulé Delor	Universitat de Barcelona	ES
Michael Louw	Lernout & Hauspie	BE
Mónica Cantero	Universitat de Barcelona	ES
Mònica Lòpez Rebollal	Universitat de Barcelona	ES
Neeme Kahusk	University of Tartu	EE
Neus Català	Universitat Politècnica de Catalunya	ES
Nicoletta Calzolari	ILC-CNR Pisa	IT
Paul Boersma	University of Amsterdam	NL
Pavel Sevecek	Masaryk University	CZ
Pedro Luis Diez-Orzas	Lernout & Hauspie	BE
Philippe Forest	Lernout & Hauspie	BE
Piek Vossen	University of Amsterdam	NL
Pierre-Francois Marteau	Bertin Technologies	F
Rita Marinelli	ILC-CNR Pisa	IT
Rudy Montigny	Lernout & Hauspie	BE
Salvador Climent Roca	Universitat de Barcelona	ES
Sílvia Calvet	Universitat de Barcelona	ES
Wim Peters	University of Sheffield	GB
Xavier Carreras	Universitat Politècnica de Catalunya	ES
Yann Fernandez	LIA, Avignon	F
Yorick Wilks	University of Sheffield	GB

1. Introduction

EuroWordNet[2] is a multilingual lexical database with wordnets for several European languages, which are structured along the same lines as the Princeton WordNet (Fellbaum 1998). WordNet contains information about nouns, verbs, adjectives and adverbs in English and is organized around the notion of a *synset*. A synset is a set of words with the same part-of-speech that can be interchanged in a certain context. For example, {car; auto; automobile; machine; motorcar} form a synset because they can be used to refer to the same concept. A synset is often further described by a gloss: "4-wheeled; usually propelled by an internal combustion engine". Finally, synsets can be related to each other by semantic relations, such as hyponymy (between specific and more general concepts), meronymy (between parts and wholes), cause, etc. as is illustrated in Figure 1.

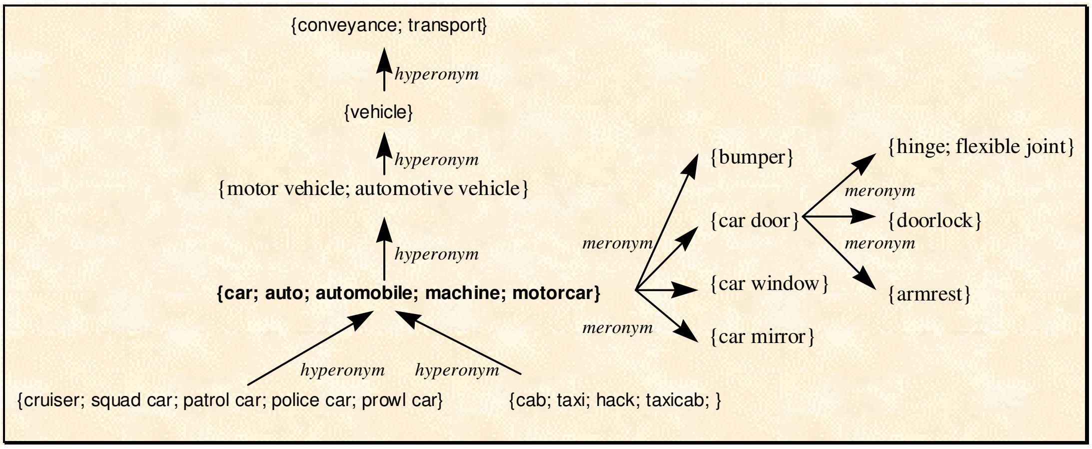


Figure 1: Synsets related to "car" *in its first sense in WordNet1.5.*

In this example, taken from WordNet1.5, the synset {car; auto; automobile; machine; motorcar} is related to:

- a more general concept or the hyperonym synset: {motor vehicle; automotive vehicle},
- more specific concepts or hyponym synsets: e.g. {cruiser; squad car; patrol car; police car; prowl car} and {cab; taxi; hack; taxicab},
- parts it is composed of: e.g. {bumper}; {car door}, {car mirror} and {car window}.

Each of these synsets is again related to other synsets as is illustrated for {motor vehicle; automotive vehicle} that is related to {vehicle}, and {car door} that is related to other parts: {hinge; flexible joint}, {armrest}, {doorlock}. By means of these and other semantic/conceptual relations, all word meanings in a language can be interconnected, constituting a huge network or wordnet. Such a wordnet can be used for making semantic inferences (what things can be used as *vehicles*), for finding alternative expressions or wordings (what words can refer to *vehicles*), or for simply expanding words to sets of semantically related or close words, in e.g. information retrieval. Furthermore, semantic networks give information on the lexicalization patterns of languages, on the conceptual density of areas of the vocabulary and on the distribution of semantic distinctions or relations over different areas of the vocabulary. In Fellbaum (1998) a detailed description is given of the history, background and characteristics of the Princeton WordNet.

Each of the European wordnets is a similar network of relations between word meanings in a specific language. The semantic relations are therefore considered as language-internal relations (see below). In addition to the language-internal relations, each synset is also linked to the closest synset in the Princeton WordNet1.5. By storing the wordnets in a central lexical database system we thus created a multilingual database, where the synsets from WordNet1.5 function as an inter-lingual index. In this database it is possible to go from one synset in a wordnet to a synset in another wordnet, which is

[2] EuroWordNet (LE2-4003 and LE-8328) is funded by the European Community within the Telematics Application Programme of the 4th Framework (DG-XIII, Luxembourg). The project started March 1996 and ended July 1999.

linked to the same WordNet1.5 concept. Such a multilingual database is useful for cross-language information retrieval, for transfer of information from one resource to another or for simply comparing the different wordnets. A comparison may tell us something about the consistency of the relations across wordnets, where differences may point to inconsistencies or to language-specific properties of the resources, or also to properties of the language itself. In this way, the database can also be seen as a powerful tool for studying lexical semantic resources and their language-specificity.

In EuroWordNet, we initially worked on 4 languages: Dutch, Italian, Spanish and English. In an extension to the project, the database was extended with German, French, Estonian and Czech. The wordnets are limited to nouns and verbs, but adjectives and adverbs are included in so far they are related to nouns and verbs (see section 2 for the relations that may hold across parts-of-speech). The vocabulary comprises all the generic and basic words of the languages: i.e. it includes all the meanings and concepts that are needed to relate more specific meanings, and all the words that occur most frequently in general corpora. For the domain of computer terminology, sub-vocabulary has been added to illustrate the possibility of integrating terminology in such a general-purpose lexicon.

The following institutes have been responsible for building the wordnets:

Dutch:	the University of Amsterdam (co-ordinator of EuroWordNet). NL.
Spanish:	the 'Fundacíon Universidad Empresa' (a co-operation of UNED Madrid, Politecnica de Catalunya in Barcelona, and the University of Barcelona). ES.
Italian:	Istituto di Linguistica Computazionale, C.N.R., Pisa. IT.
English:	University of Sheffield (adapting the English wordnet). GB.
French:	Université d' Avignon and Memodata at Avignon. F.
German:	Universität Tübingen. DE.
Czech:	University of Masaryk at Brno. CZ.
Estonian:	University of Tartu, EE.

Each of these institutes was responsible for the construction of their national wordnet, where most of them used material and resources developed outside the project (among which lexical resources from the publishers Van Dale for Dutch and Bibliograf for Spanish). The task of Sheffield has been different because of the existence of WordNet for English. Their role consisted of adapting the Princeton WordNet for the changes made in EuroWordNet and controlling the interlingua that connects the wordnets.

In addition to the wordnet builders there have been 3 industrial users in the project:

- Bertin & Cie, Plaisir, France
- Xerox Research Centre, Meylan, France
- Novell Linguistic Development (changed to Lernout & Hauspie during the project), Antwerp, Belgium

They demonstrated the use of the database in their (multilingual) information-retrieval applications. Novell also had an additional role as the developer of the central EuroWordNet database Polaris and the database viewer Periscope.

On a longer term we expect that EuroWordNet will open up a whole range of new applications and services in Europe at a trans-national and trans-cultural level. It will give information on the typical lexicalization patterns across languages, which will be crucial for machine translation and language learning systems. It will give non-native users and non-skilled writers the possibility to navigate or browse through the vocabulary of a language in new ways, giving them an overview of expression which is not feasible in traditional alphabetically-organized resources. Finally, it will stimulate the development of sophisticated lexical knowledge bases that are crucial for a whole gamut of future applications, ranging from basic information retrieval to question/answering systems, language understanding and expert systems, from summarizers to automatic translation tools and resources.

In this document, we will give a general description of the database. The 4 main sections cover the design of the database (section 2), the general methodology (section 3), the main database functionality (section 4) and the content of the CD-rom (section 5), respectively. In addition to this general document, there wil be separate documents that describe the content of the each wordnet and a comparison across each set of wordnets:

- EuroWordNet-1: A comparison of the Dutch, English, Italian and Spanish wordnets
- EuroWordNet-2: A comparison of the German, French, Czech and Estonian wordnets

In the individual wordnet documents information is given on the size and quantity of the data, as well more specific details on the methods of building. The comparison consists of an overview of the quantitive properties of the wordnets and their compatibility measured in terms of the equivalences to which they are linked. These documents are released with the databases. All this information on EuroWordNet and more can also be downloaded from http://www.hum/uva.nl/~ewn.

The next section on the database design will give an overview of the different modules (section 2.1), the language internal structures (section 2.2), the multilingual structure (section 2.3), the word sense or synset variant structure (section 2.4), and an explanation of the plain text representation of the data (section 2.5).

2. Design of the multilingual database

The design of the EuroWordNet-database is first of all based on the structure of the Princeton WordNet and specifically version WordNet1.5. The notion of a synset and the main semantic relations have been taken over in EuroWordNet. However, some specific changes have been made to the design of the database, which are mainly motivated by the following objectives:

1) to create a multilingual database;
2) to maintain language-specific relations in the wordnets;
3) to achieve maximal compatibility across the different resources;
4) to build the wordnets relatively independently (re)-using existing resources;

The most important difference of EuroWordNet with respect to WordNet is its multilinguality, which however also raises some fundamental questions with respect to the status of the monolingual information in the wordnets. In principle, multilinguality is achieved by adding an equivalence relation for each synset in a language to the closest synset in WordNet1.5. Synsets linked to the same WordNet1.5 synset are supposed to be equivalent or close in meaning and can then be compared. However, what should be done with differences across the wordnets? If 'equivalent' words are related in different ways in the different resources, we have to make a decision about the legitimacy of these differences. For example, in the Dutch wordnet we see that *hond* (dog) is both classified as *huisdier* (pet) and *zoogdier* (mammal). However, there is no equivalent for *pet* in Italian, and the Italian *cane*, which is linked to the same synset *dog*, is only classified as a *mammal* in the Italian wordnet.

In EuroWordNet, we take the position that it must be possible to reflect such differences in lexical semantic relations. The wordnets are seen as linguistic ontologies rather than ontologies for making inferences only. In an inference-based ontology it may be the case that a particular level or structuring is required to achieve a better control or performance, or a more compact and coherent structure. For this purpose it may be necessary to introduce artificial levels for concepts which are not lexicalized in a language (e.g. *natural object*, *external body parts*), or it may be necessary to neglect levels (e.g. *watchdog*) that are lexicalized but not relevant for the purpose of the ontology. A linguistic ontology, on the other hand, exactly reflects the lexicalization and the relations between the words in a language. It is a "wordnet" in the true sense of the word and therefore captures valuable information about conceptualizations that are lexicalized in a language: what is the available fund of words and expressions in a language. In addition to the theoretical motivation there is also a practical motivation for considering the wordnets as autonomous networks. To be more cost-effective, they have (as far as possible) been derived from existing resources, databases and tools. Each sites therefore had a different starting point for building their local wordnet, making it necessary to allow for a maximum of flexibility in producing the wordnets and structures.

2.1. The Database Modules

To be able to maintain the language-specific structures and to allow for the separate development of independent resources, we make a distinction between the language-specific modules and a separate language-independent module. Each language module represents an autonomous and unique language-specific system of language-internal relations between synsets. Equivalence relations between the synsets in different languages and WordNet1.5 are made explicit in the so-called Inter-Lingual-Index (ILI). Each synset in the monolingual wordnets has at least one equivalence relation with a record in this ILI, either directly or indirectly via other related synsets. Language-specific synsets linked to the same ILI-record should thus be equivalent across the languages, as is illustrated in Figure 2 for the language-specific synsets linked to the ILI-record *drive*.

Figure 2 further gives a schematic presentation of the different modules and their inter-relations. In the middle, the language-external modules are given: the ILI, a Domain Ontology and a Top Concept Ontology. The ILI consists of a list of so-called ILI-records (ILIRs) which are related to word-meanings in the language-internal modules, (possibly) to one or more Top Concepts and (possibly) to domains. The language-internal modules then consist of a lexical-item-table indexed to a set of word-meanings, between which the language-internal relations are expressed.

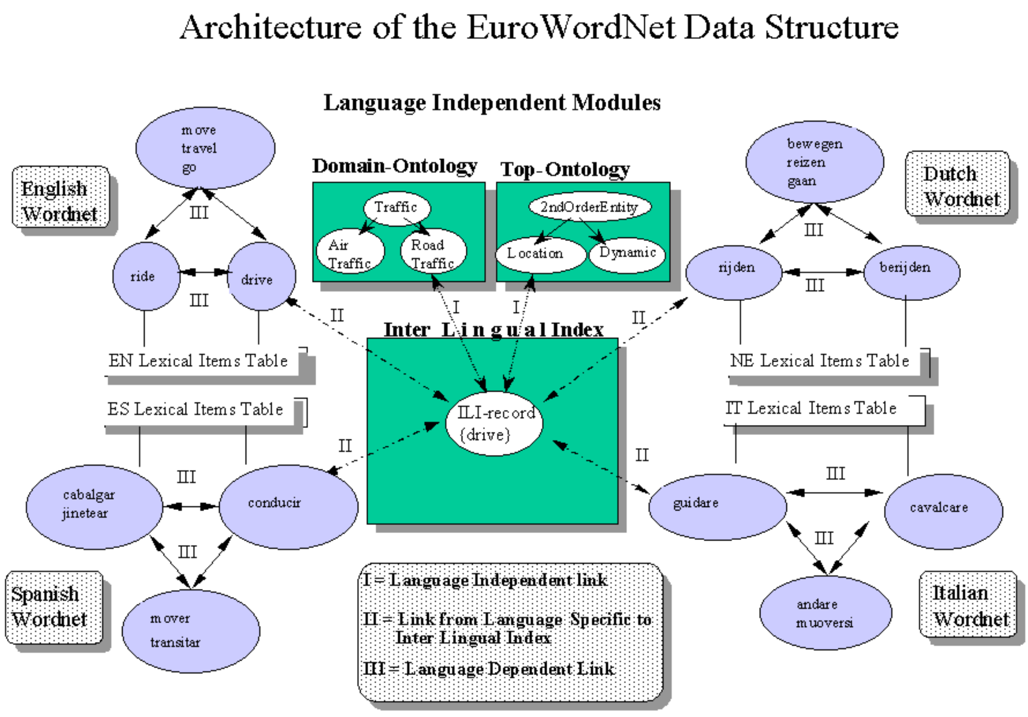


Figure 2. The global architecture of the EuroWordNet database.

The ILI is an unstructured list of meanings, mainly taken from WordNet1.5, where each ILI-record consists of a synset, an English gloss specifying the meaning and a reference to its source. The only purpose of the ILI is to mediate between the synsets of the language-specific wordnets. No relations are therefore maintained between the ILI-records as such. The development of a complete language-neutral ontology is considered to be too complex and time-consuming given the limitations of the project. As an unstructured list, there is no need to discuss changes or updates to the index from a many-to-many perspective. Note that it will nevertheless be possible to indirectly see a structuring of a set of ILI-records by viewing the language-internal relations of the language-specific concepts that are related to the set of ILI-records. Since WordNet1.5 is linked to the index in the same way as any of the other wordnets, it is still possible to recover the original internal organization of the synsets in terms of the semantic relations in WordNet1.5.

The advantages of an interlingua such as the Inter-Lingual-Index are well-known in MT translation (Copeland et al. 1991, Nirenburg 1989):

1. it is not necessary to specify many-to-many equivalence relations between each language-pair and to have consensus across all the groups on the equivalence relations: each group only considers the equivalence relations to the Index.
2. new languages can be added without having to reconsider the equivalence relations for the other languages.
3. it is possible to adapt the Inter-Lingual-Index as a central resource to make the matching more efficient or precise.

In section 2.3, we will describe how we adapted the ILI to provide a more efficient mapping across the wordnets. Updates can be made relatively easy because the ILI lacks any further structure.

Some language-independent structuring of the ILI is nevertheless provided by two separate ontologies, which may be linked to ILI records:

- the Top Concept ontology, which is a hierarchy of language-independent concepts, reflecting important semantic distinctions, e.g. Object and Substance, Location, Dynamic and Static;

- a hierarchy of domain labels, which are knowledge structures grouping meanings in terms of topics or scripts, e.g. Traffic, Road-Traffic, Air-Traffic, Sports, Hospital, Restaurant;

Both the Top Concepts and the domain labels can be transferred via the equivalence relations of the ILI-records to the language-specific meanings, as is illustrated in Figure 2. The Top Concepts *Location* and *Dynamic* are for example directly linked to the ILI-record *drive* and therefore indirectly also apply to all language-specific concepts related to this ILI-record. Via the language-internal relations, the Top Concept can be further inherited by all other related language-specific concepts. The main purpose of the Top Ontology is to provide a common framework for the most important concepts in all the wordnets. It consists of 63 basic semantic distinctions that classify a set of 1300 ILI-records representing the most important concepts in the different wordnets. The classification has been verified by the different sites, so that it holds for all the language-specific wordnets. In section 3.4, we will further describe the Top Ontology and its motivation.

The domain-labels can be used directly in information retrieval (and also in language-learning tools and dictionary publishing) to group concepts in a different way, based on scripts rather than classification. Domains can also be used to separate the generic from the domain-specific vocabularies. This is important to control the ambiguity problem in Natural Language Processing. So far we have only included domain labels for computer terminology in EuroWordNet. However, users of the database can freely add domain labels to the ILI or adjust the top ontology without having to access or consider the language-internal relations of each wordnet. In the same way, it is possible to extend the database with other ontologies provided that they are specified according to the EuroWordNet format and include a proper linking to the ILI.

Once the wordnets are properly linked to the ILI, the EuroWordNet database makes it possible to compare wordnet fragments via the ILI and to track down differences in lexicalization and in the language-internal relations. This is illustrated in Figure 3, which is taken from the graphical interface to the EuroWordNet database, called Periscope (Cuypers and Adriaens 1997). The top-half of the screen-dump shows a window with a fragment of the Dutch wordnet at the left and a similar fragment of WordNet1.5 at the right. The bottom window shows a similar parallel view for the Italian and Spanish wordnets. Each synset in these windows is represented by a rectangular box followed by the synset members. On the next line, the closest Inter-Lingual-Index concept is given, following the = sign (which indicates direct equivalence). In this view, the ILI-records are represented by an English gloss. Below a synset-ILI pair, the language-internal relations can be expanded, as is done here for the hyperonyms. The target of each relation is again represented as a synset with the nearest ILI-equivalent (if present). The first line of each wordnet gives the equivalent of *cello* in the 4 wordnets. In this case, they are all linked to the same ILI-record, which indirectly suggests that they should be equivalent across the wordnets as well. We also see that the hyperonyms of *cello* are also equivalent in the two windows, as is indicated by the lines connecting the ILI-records. Apparently, the structures are parallel across the Dutch wordnet and WordNet1.5 on the one hand and the Spanish and Italian wordnets on the other. However, we see that the intermediate levels for *bowed stringed instrument* and *stringed instrument* in the Dutch wordnet and WordNet1.5 are missing both in Italian and Spanish. Had we compared other wordnet pairs, the intermediate synsets would be unmatched across the wordnets.

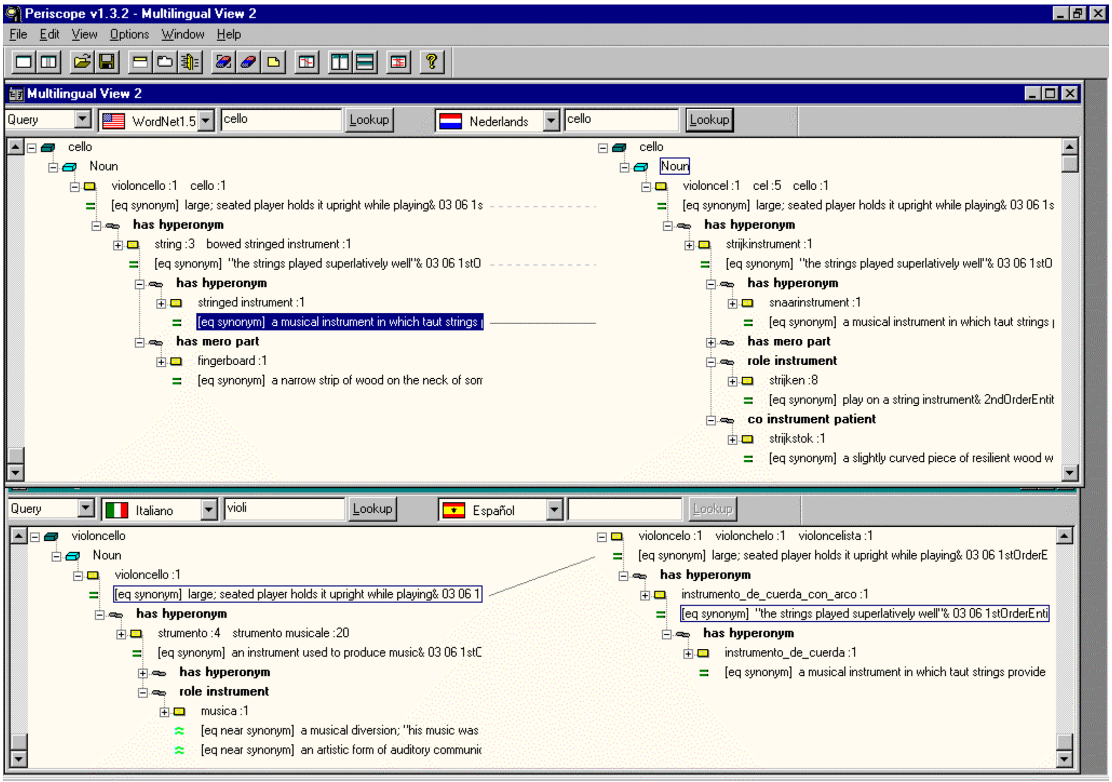


Figure 3: Parallel wordnet structures in EuroWordNet linked to the same ILI-records.

A further discussion on the advantages and disadvantages of different multilingual designs and the ways of comparing the wordnets is given in Peters et al. (1998).

Summarizing, the modular multilingual design of the EWN-database has the following advantages:

- it will be possible to use the database for multilingual information retrieval, by expanding words in one language to related words in another language via the ILI;
- the different wordnets can be compared and checked cross-linguistically which will make them more compatible;
- language-dependent differences can be maintained in the individual wordnets;
- it will be possible to develop the wordnets at different sites relatively independently;
- language-independent information such as the glosses, the domain-knowledge and the analytic Top Concepts can be stored only once and can be made available to all the language-specific modules via the inter-lingual relations;
- the database can be tailored to a user's needs by modifying the Top Concepts, the domain labels or instances, (e.g. by adding semantic features) without having to access the language-specific wordnets;

2.2. The Language Internal Relations

The EWN database is a 'relational' database in which the meaning of each word is basically described by means of its relations to other word meanings. Most of the WordNet1.5 relations, commonly accepted in various approaches to semantics, have been taken over in EWN. Nevertheless, some changes have been made with respect to WordNet1.5:

1. the use of labels to relations that make the semantic entailments more explicit and precise (e.g. conjunction of relations: a *knife* is either a *weapon* or a *piece of cutlery*, a *spoon* is both a *container* and a *piece of cutlery*);
2. the introduction of cross part-of-speech relations, so that different surface realizations of similar concepts within and across languages can still be matched (e.g. between the verb *adorn* and the noun *adornment* or the noun *death* and the adjective *dead*);

3. the addition of some extra relations to differentiate certain shallow hierarchies (e.g. semantic role relations between nouns and verbs, such as agent (*teacher*), patient (*student*), location (*school*) related to *teach*);

A crucial difference here are the relations across part-of-speech. Whereas the Princeton WordNet maintains a strict division between the different parts-of-speech, many relations between different part-of-speech are allowed in EuroWordNet. Instead of the part-of-speech distinction, EuroWordNet makes a fundamental difference between 3 types of entities following Lyons (1977):

1stOrderEntity
: Any concrete entity (publicly) perceivable by the senses and located at any point in time, in a three-dimensional space, e.g. object, substance, animal, plant, man, woman, instrument.

2ndOrderEntity
: Any Static Situation (property, relation) or Dynamic Situation, which cannot be grasped, heart, seen, felt as an independent physical thing. They can be located in time and occur or take place rather than exist; e.g. be, happen, cause, move, continue, occur, apply.

3rdOrderEntity
: Any unobservable proposition that exists independently of time and space. They can be true or false rather than real. They can be asserted or denied, remembered or forgotten. E.g. idea, thought, information, theory, plan, intention.

We will see that certain relations can only hold between certain types of entities, but that these entities can be named often by words with different parts-of-speech. The tests that are used to verify the relations are then rephrased to fit the different parts of speech but the conditions are formulated for entity types. In section 3.4, we will further describe the ontological status of these 3 types of entities.

EuroWordNet represents a more general semantic *model* that incorporates different types of important semantic relations that are extractable from dictionaries (and other sources) and of usage for NLP applications. The definition of such a broad model does not, however, imply that all possible relations for all meanings have been provided. Given the project's limitations in time and budget, the encoding of additional semantic relations has been restricted to those meanings that can be (semi-)automatically derived from our sources or to those meanings that cannot be related properly by means of the more basic relations only.

This section is further organized as follows. First, we illustrate the kind of criteria and principles we used to verify a relation between synsets (subsection 2.2.1.). In section 2.2.2., we describe the relation labels and in section 2.2.3., the different types of relations.

2.2.1. Criteria for the identification of relations between synsets

Following Cruse (1986), we created substitution tests or diagnostic frames to verify relations between synsets. Inserting two words in the test sentences will mostly evoke a strong 'normality'/'abnormality' judgement, on the basis of which the relation can be determined. For instance, synsets are identified on the basis of the possibility of a word being replaced by another in a specific context. This can be verified by the possibility of being mutually substitutable in sentence (a) for nouns, and sentence (b) for verbs:

a. X is a $Noun_1$ therefore X is a $Noun_2$
b. Y $Verb(\text{-}phrase)_1$ therefore Y $Verb(\text{-}phrase)_2$

For instance, *fiddle* and *violin* are synonyms on the basis of the 'normality' of (1a) and (1b), while *dog* and *animal* are not, due to the 'abnormality' of (2b); in a similar way, *enter* and *go into* are synonyms, while *walk* and *move* are not:

1a. It is a *fiddle* therefore it is a *violin*.
2a. It is a *violin* therefore it is a *fiddle*.
3a. It is a *dog* therefore it is an *animal*.
4a. *It is an *animal* therefore it is a *dog*.[3]

[3] '*' is used, here and in the following examples, to indicate 'semantic abnormality'.

1b. John *entered* the room therefore John *went into* the room.
2b. John *went into* the room therefore John *entered* the room.
3b. The dog *walked* therefore the dog *moved*.
4b. *The dog *moved* therefore the dog *walked*.

Similar tests have been developed for every relation in EWN, in each of the different languages. Note that these tests are devised to detect semantic relations only and are not intended to cover differences in register, style or dialect between words. The tests not only provide us with a common definition for carrying out the work independently but can also be used by external people to verify the quality of our work. In Alonge (1996) and Climent et al. (1996), tests are described for most relations in English, Dutch, Spanish and Italian. These documents can be downloaded from the EuroWordNet WWW-site http://www.hum.uva.nl/~ewn. Below, we will only give the English tests to illustrate the meaning and use of each relation.

In addition to the tests there are some other principles which can be used for encoding the relations. One of them is the *Economy principle* (Dik, 1978) which states that a word should not be defined in terms of more general words when there are more specific words that can do the job. If we apply this to hyperonymy/hyponymy[4] the principle can be formalized as follows:

> *If a word W_1 is the hyperonym of W_2 and W_2 is the hyperonym of W_3 then W_3 should not be directly linked to W_1 but to W_2.*[5]

This principle should prevent intermediate levels from being skipped, i.e. senses from being (directly) linked too high up in the hierarchy.

A second principle is the *Compatibility principle,* which can be formulated as:

> *If a word W_1 is related to W_2 via relation R_1, W_1 and W_2 cannot be related via relation R_n, where R_n is defined as a relation distinct from R_1.*

In other words, if two word senses are linked by a particular type of relation (e.g. as synonyms), then they cannot be linked by means of any other relation (e.g. as antonyms). Although this general rule directly follows from the way in which the relations are defined, there are cases in which it is somehow difficult to maintain it. For instance, group nouns or collectives, such as *cutlery* and *furniture,* can easily be linked by *hyponymy* and *meronymy* to the terms representing individual items included in the groups, such as *fork* and *table* respectively. Some relations will then have priority over other relations, in the above case hyponymy over meronymy (cf. Vossen et al., (1998), for a more detailed discussion).

Finally, we have provided in some cases more specific tests in addition to more general tests. This is done because the more specific tests yield stronger intuitions on the validity of relations. It is easier to agree with a specific test than with a more general abstract test. If the specific test fails or is questionable, it is still possible to use the more general test.

2.2.2. Relation Labels

A major difference between the EWN database and the structure of WN1.5 is the possibility of adding labels to the relations. These labels are needed to differentiate the precise semantic implications that follow from the defined relations. The following types of labels have been distinguished:

- conjunction or disjunction of multiple relations of the same type related to a synset;
- (non-)factivity of causal relations;
- reversal of relations;
- negation of relations.

---

[4] What we indicate here as *hyperonymy* is sometimes spelled as *hypernymy* (e.g., in WN). Moreover, in WN a distinction is drawn between hyperonymy (the relation occurring between nouns) and *troponymy* (occurring between verbs), because of the different nature of the relation linking verbs to their superordinates discussed in Fellbaum (1990) (but cf. also Cruse 1986). Although we generally agree with Fellbaum's remarks on this issue, we have decided to use the traditional label *hyperonymy* also for the relation linking verbs.

[5] Of course, since the hyponymy (or IS-A) relation is a *transitive* relation, W_3 will be a sub-hyponym of W_1.

2.2.2.1. Conjunction /Disjunction

The *conjunction* and *disjunction* labels are used to explicitly mark the status of multiple relations of the same type displayed by a synset. In WN1.5 the interpretation is not explicit. It is a matter of practice that e.g. multiple meronyms linked to the same synset are automatically taken as conjunctives: “all the parts together constitute the holonym *car*”. Furthermore, we see that different senses are distinguished for words referring to parts belonging to different kinds of holonyms (e.g. *door*):

door1 -- (a swinging or sliding barrier that will close the entrance to a room or building; “he knocked on the door”; “he slammed the door as he left”) PART OF: doorway, door, entree, entry, portal, room access
door 6 -- (a swinging or sliding barrier that will close off access into a car; “she forgot to lock the doors of her car”) PART OF: car, auto, automobile, machine, motorcar.

In more traditional resources, similar relations are expressed often by explicit disjunction or conjunction of words in the same definition. Note that this is also done in the definition of the first sense of *door* in WN1.5 where *room* and *building* are coordinated in the gloss. In EWN, disjunction and conjunction can be indicated explicitly by a relation label or feature:

{airplane}			{door}	
HAS_MERONYM: c1	{door}		HAS_HOLONYM: d1	{car}
HAS_MERONYM: c2d1	{jet engine}		HAS_HOLONYM: d2	{room}
HAS_MERONYM: c2d2	{propeller}		HAS_HOLONYM: d3	{airplane}

Here c1, c2 and d1, d2, d3 represent conjunction and disjunction respectively, where the index keeps track of the scope of nested combinations. For example, in the case of *airplane* we see that either a *propeller* or a *jet engine* constitutes a part that is combined as the second constituent with *door*. Note that one direction of a relationship can have a conjunctive index, while the reverse can have a disjunctive one. Finally, when conjunction and disjunction labels are absent, multiple relations of the same type are interpreted as non-exclusive disjunction (and/or).

Conjunction and *disjunction* may also apply to other relations than meronymy such as hyponymy: a *spoon* is both a *container* and a *piece of cutlery* at the same time. In other cases, hyperonyms are clearly disjunctive: an *albino* either is an *animal*, *human* or a *plant*, a *threat* may be a *person*, *idea* or *thing*.

2.2.2.2. Factivity

Lyons (1977) distinguishes different types of causality on the basis of the factivity of the effect:

- factive: event E_1 implies the causation of E_2

  “to kill causes to die”

- non-factive: E_1 probably or likely causes event E_2 or E_1 is intended to cause some event E_2

  “to search may cause to find”.

The label *non-factive* is added to a causal relation to indicate that the relation does not necessarily hold. Absence of a label indicates factivity by default.

2.2.2.3. Reversed

It is a requirement of the database that every relation has a reverse counter-part. However, there are relations that are conceptually bi-directional, and others that are not. In the case of hyperonymy/hyponymy, the relation holds in both directions: e.g. since *hammer* is a hyponym of *hand tool*, *hand tool* is a hyperonym of *hammer*. In the case of, for example, a meronymy relation the implicational direction may, instead, vary:

hand	HAS_MERONYM	finger	
finger	HAS_HOLONYM	hand	
car	HAS_MERONYM	door	
door	HAS_HOLONYM	car	*reversed*
computer	HAS_MERONYM	disk drive	*reversed*
disk drive	HAS_HOLONYM	computer	

In the case of *finger* and *hand* the dependency or implication holds in both directions. In the case of *car* and *door* however, we see that *car* always implies the meronym *door* but *door* does not necessarily imply the holonym *car*. For *computer* and *disk drive*, we see the opposite dependency: a *disk drive* is a part of a *computer* but not every *computer* has a *disk drive*. Since relations that are stated in one direction are automatically reversed in the database, it is not possible to distinguish these different directions of implication, unless they are labelled. Therefore, the label *reversed* is added to those relations that are not necessarily implied or not conceptually salient but are only the result of the automatic reversal.[6]

2.2.2.4. Negation

The negation label *negative* explicitly expresses that a relation does not hold:

macaque	HAS_MERONYM	tail	
Barbary ape	HAS_MERONYM	tail	*negative*

Such a label can be used to explicitly block certain implications. For instance, a *macaque* has a tail. Normally, parts are inherited along a taxonomy, thus, being a kind of macaque, the *Barbary ape* should have a tail. However, a *Barbary ape* does not have a tail, and by using the label *negative* this inference can be blocked. In the following subsections, more examples will be given of the use of these labels when discussing relations.

2.2.3. The subtypes of language-internal relations

The most important relation in WN1.5 is synonymy, which is implicit in the notion of a synset. The other relations encoded in WN1.5 are given in Table 1 together with examples for the various parts-of-speech (POS) linked:

Table 1: WordNet1.5 Relations

Relation	*PoS linked*	*Example*	*EWN*
ANTONYMY	noun/noun; verb/verb; adjective/adjective	man/woman; enter/exit; beautiful/ugly	yes
HYPONYMY	noun/noun	slicer/knife	yes
MERONYMY	noun/noun	head/nose	yes
ENTAILMENT	verb/verb	buy/pay	**SUBEVENT** or **CAUSE**
TROPONYM	verb/verb	walk/move	**HYPONYMY**
CAUSE	verb/verb	kill/die	yes
ALSO SEE	verb/adjective		no
DERIVED FROM	adjective/adverb	beautiful/beautifully	yes
ANTONYM	noun/noun; verb/verb	heavy/light	yes
ATTRIBUTE	noun/adjective	size/small	**XPOS_HYPONYM**
RELATIONAL ADJ	adjective/noun	atomic/ atomic bomb	**PERTAINS TO**
SIMILAR TO	adjective/adjective	ponderous/heavy	no
PARTICIPLE	adjective/verb	elapsed/ elapse	no

[6] Currently, if a new wordnet is imported in the database, a relation is expressed in one direction from the source concept to the target concept. The database will first automatically generate the corresponding reversed relation, adding the label reversed. Only if the relation is also explicitly expressed in the other direction, the database will remove the reverse label when resolving the relations. It is also possible to explicitly specify labels in the import file. The database will honour these specification.

The last column indicates what relations have been taken over in EuroWordNet or have been converted to other relations.

The next two tables then give the complete list of Language-Internal-Relations in EuroWordNet. The first table gives the relations between synsets, and the second table between other data types (instances and variants or synset members). For each relation the following information is given:

i) its name,
ii) the parts of speech linked (with an indication of the 'direction' of the linking: < or >),
iii) further relation labels that may apply,
iv) the type of data linked (i.e. synsets, synset variants or instances).

The part-of-speech constraints are the formal constraints that will be checked by the EuroWordNet database, Polaris. This is because the part-of-speech is more easily verifiable than the differentiation between different entity types. Nevertheless, underlying many limitations between the part-of-speech combinations are still constraints on the types of entities, e.g. a CAUSE relation can only have a 2ndOrderEntity as a target (which can be realized as a noun, verb or adjective/adverb in the current set of languages).

Parts of Speech:
- N = noun
- V = verb
- AdjAdv = Adjective or Adverb
- PN = pronoun or name

Labels:
- dis = disjunctive
- con = conjunctive
- rev = reversed
- non-f = non-factive
- neg = negative

Data types:
- Syn = synset
- I = instance
- VA = synset variant

Table 2: Language Internal Relations between synsets in EuroWordNet

Relation Type	Parts of Speech	Labels	Data Types
NEAR_SYNONYM	N<>N, V<>V		Syn <>Syn
XPOS_NEAR_SYNONYM	N<>V, N<>AdjAdv, V<>AdjAdv		Syn <>Syn
HAS_HYPERONYM	N>N, V>V	dis, con	Syn <>Syn
HAS_HYPONYM	N>N, V>V	dis	Syn <>Syn
HAS_XPOS_HYPERONYM	N>V, N>AdjAdv, V>AdjAdv, V>N, AdjAdv>N, AdjAdv>V	dis, con	Syn <>Syn
HAS_XPOS_HYPONYM	N>V, N>AdjAdv, V>AdjAdv, V>N, AdjAdv>N, AdjAdv>V	dis	Syn <>Syn
HAS_HOLONYM	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_PART	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_MEMBER	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_PORTION	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_MADEOF	N>N	dis, con, rev, neg	Syn <>Syn
HAS_HOLO_LOCATION	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERONYM	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_PART	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_MEMBER	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_MADEOF	N>N	dis, con, rev, neg	Syn <>Syn
HAS_MERO_LOCATION	N>N	dis, con, rev, neg	Syn <>Syn
ANTONYM	N<>N, V<>V		Syn <>Syn
NEAR_ANTONYM	N<>N, V<>V		Syn <>Syn
XPOS_NEAR_ANTONYM	N<>V, N<>AdjAdv, V<>AdjAdv		Syn <>Syn
CAUSES	V>V, N>V, N>N, V>N, V>AdjAdv, N>AdjAdv	dis, con, non-f, rev , neg	Syn <>Syn
IS_CAUSED_BY	V>V, N>V, N>N, V>N, AdjAdv>V, AdjAdv>N	dis, con, non-f, rev, neg	Syn <>Syn
HAS_SUBEVENT	V>V, N>V, N>N, V>N	dis, con, rev, neg	Syn <>Syn
IS_SUBEVENT_OF	V>V, N>V, N>N, V>N	dis, con, rev, neg	Syn <>Syn
ROLE	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_AGENT	N>V, N>N	dis, con, rev, neg	Syn <>Syn
ROLE_INSTRUMENT	N>V, N>N	dis, con, rev, neg	Syn <>Syn
ROLE_PATIENT	N>V, N>N	dis, con, rev, neg	Syn <>Syn
ROLE_LOCATION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_DIRECTION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_SOURCE_DIRECTION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_TARGET_DIRECTION	N>V, N>N, AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
ROLE_RESULT	N>V, N>N	dis, con, rev, neg	Syn <>Syn
ROLE_MANNER	AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
INVOLVED	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_AGENT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
INVOLVED_PATIENT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
INVOLVED_INSTRUMENT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
INVOLVED_LOCATION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_DIRECTION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_SOURCE_DIRECTION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_TARGET_DIRECTION	V>N, N>N, V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
INVOLVED_RESULT	V>N, N>N	dis, con, rev, neg	Syn <>Syn
CO_ROLE	N>N	rev	Syn <>Syn
CO_AGENT_PATIENT	N>N	rev	Syn <>Syn
CO_AGENT_INSTRUMENT	N>N	rev	Syn <>Syn
CO_AGENT_RESULT	N>N	rev	Syn <>Syn
CO_PATIENT_AGENT	N>N	rev	Syn <>Syn
CO_PATIENT_INSTRUMENT	N>N	rev	Syn <>Syn
CO_PATIENT_RESULT	N>N	rev	Syn <>Syn
CO_INSTRUMENT_AGENT	N>N	rev	Syn <>Syn
CO_INSTRUMENT_ PATIENT	N>N	rev	Syn <>Syn
CO_INSTRUMENT_RESULT	N>N	rev	Syn <>Syn
CO_RESULT_AGENT	N>N	rev	Syn <>Syn
CO_RESULT_PATIENT	N>N	rev	Syn <>Syn
CO_RESULT_INSTRUMENT	N>N	rev	Syn <>Syn
IN_MANNER	V>AdjAdv, N>AdjAdv	dis, con, rev, neg	Syn <>Syn
MANNER_OF	AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
BE_IN_STATE	N>AdjAdv, V>AdjAdv	dis, con, rev, neg	Syn <>Syn
STATE_OF	AdjAdv>N, AdjAdv>V	dis, con, rev, neg	Syn <>Syn
FUZZYNYM	N<>N, V<>V		Syn <>Syn
XPOS_FUZZYNYM	N<>V, V<>AdjAdv, N<>AdjAdv		Syn <>Syn

Table 3: Language-Internal Relations between other data types in EuroWordNet

Relation Type	Parts of Speech	Labels	Data Types
IS_DERIVED_FROM	N, V, AdjAdv (across all)		VA<>VA
HAS_DERIVED	N, V, AdjAdv (across all)		VA<>VA
DERIVATION	N, V, AdjAdv (across all)		VA<>VA
ANTONYM	N<>N, V<>V, AdjAdv <> AdjAdv		VA<>VA
PERTAINS_TO	AdjAdv>N, AdjAdv>V		VA<>VA
IS_PERTAINED_TO	N>AdjAdv, V>AdjAdv		VA<>VA
HAS_INSTANCE	N>PN		Syn>I
BELONGS_TO_CLASS	PN>N		I>Syn

In the next subsections we will discuss each relation and give some examples.

2.2.3.1. Synonymy

Synonymy is the basis for the organization of the database in synsets. In principle all semantically equivalent words should belong to the same synsets (where they can be differentiated by labels on the appropriate usage). A formal definition of synonymy, given by Leibniz, is:

> *"two expressions are synonyms if the substitution of one for the other* ***never*** *change the truth value of a sentence in which the substitution is made"*

However, true synonyms are rarely found in language. Miller and Fellbaum (1990) therefore suggest to use a weaker notion of synonymy, namely **'semantic similarity'**, which is defined as:

> *"two expressions are synonymous in* ***a*** *linguistic context C if the substitution of one for the other in C does not alter the truth value"* (Miller et al., 1990).

One such context is thus already sufficient to allow a synonymy relation between word meanings. This leaves room for different interpretations. Following Miller and Fellbaum (1990) and Cruse (1986), what seems clear is however that synonymy should be a symmetric relation, that is, if X is 'semantically similar' to Y, then Y is equally 'semantically similar' to X, while, obviously, hypernymy-hyponymy should be asymmetric.

In EuroWordNet, we further mean by **semantically-equivalent** that two words denote the same range of entities, irrespective of the morpho-syntactic differences, differences in register, style or dialect or differences in pragmatic use of the words. Another, more practical, criterion which follows from the above homogeneity principle is that two words which are synonymous cannot be related by any of the other semantic relations defined. This would mean that, for example, the following variants belong to the same synset:

> {people, folks}
> {cop, pig, policeman, police officer}

but it also means that "person" and "police force" cannot belong to these synsets because there is another semantic relation: "**member-group**" that can be used to relate them (even though they are in many cases interchangeable in language use).

Strictly speaking, this definition allows for synonymy across parts-of-speech, e.g. "shot N", "shoot V". However, since the distinction between part-of-speech (as an intrinsic property of WordNet1.5) is crucial to many systems using WordNet1.5 we have decided to use a separate relation for synonymy (and also hyponymy) across parts-of-speech: XPOS_NEAR_SYNONYM (see below)

The above claims can be formulated as follows for nouns and verbs:

- in any sentence S where Noun1 is the head of an NP which is used to identify an entity in discourse another noun Noun2 which is a synonym of Noun1 can be used as the head of the same NP without resulting in semantic anomaly. And vice versa for Noun2 and Noun1.
- in any sentence S where $Verb_1$ is the head of a VP which is used to identify a situation in discourse another verb $Verb_2$, which is a synonym of $Verb_1$, can be used as the head of the same VP without resulting in semantic anomaly. And vice versa for Verb2 and Verb1.

From this we can derive the following tests for synonymy between nouns and verbs respectively:

Test 1		**Synonymy between nouns**
yes	a	if it is (a/an) X then it is also (a/an) Y
yes	b	if it is (a/an) Y then it is also (a/an) X
Conditions:		X and Y are singular or plural nouns
Example:	a	if it is a fiddle then it is a violin
	b	if it is a violin then it is a fiddle
Effect:		synset variants {fiddle, violin}

Test 2		**Synonymy between verbs**
yes a		If something/someone/it Xs then something/someone/it Ys
yes b		If something/someone/it Ys then something/someone/it Xs
Conditions:		- X is a verb in the third person singular form - Y is a verb in the third person singular form - there are no specifying PPs that apply to the X-phrase or the Y-phrase
Example:	a	If something/someone/it begins then something/someone/it starts
	b	If something/someone/it starts then something/someone/it begins
Effect:		synset variants: {begin, start}

The substitution sentences for synonymy are the same as for hyponymy, with the only difference that synonyms are mutually exclusive whereas words with a hyponymy relation are partially interchangeable (see below).

In many cases there is a close relation between words but not sufficient to make them members of the same synset, i.e.: they do not yield clear scores for the previous test or their hyponyms cannot be interchanged. For these cases we can use the NEAR_SYNONYM relation. The next test expresses differences in the range of hyponyms across close concepts:

Test 3		**Near_synonymy between nouns that differ in range of hyponyms.**		
yes a		if it is a/an X then it is also a kind of Y but you usually do not call Z_n Ys		
yes b		if it is a/an Y then it is also a kind of X but you usually do not call Z_m Xs		
Conditions:		Z_n are hyponyms of X, Z_m are hypnyms of Y.		
Example:	a	if it is a tool then it is also an instrument but you usually do not call hammers, screw drivers, etc. instruments		
	b	if it is an instrument then it is also a tool but you usually do not call measure intruments, musical instruments , etc. tools		
Effect:		tools	NEAR_SYNONYM	instrument
		instrument	NEAR_SYNONYM	tools

Using the NEAR_SYNONYMY relation we can keep sets of hyponyms separate while we can still encode that two synsets are closer in meaning than other co-hyponyms, e.g. *tool* versus *body*, *instrument* versus *fruit* which are all subtypes of *object*.

We mentioned that WordNet1.5 maintains a strict separation between the different parts-of-speech, but in EuroWordNet explicit relations across parts-of-speech may occur. The first relation to be discussed is synonymy across part-of-speech, as between “move” and “movement”. The POS difference leads to subtle differences in meaning (such as argument reduction of nominalizations) but in many cases languages offer a choice between a noun, verb or adjective to name the same situation or event. Even stronger, there are many cases of part-of-speech mismatch across languages, which can only be translated by different morpho-syntactic realizations.

Cross-part-of-speech relations are often derivational, but very different meanings can be associated with these derivations, e.g. the noun *cut* can both be the event or the result of the event. Since this information is not always predictable it is useful to make the relation explicit. In this subsection, we will discuss near-synonymy relations across part-of-speech. Later we will also describe cross-pos hyponymy, antonymy and causal relations across parts of speech. In all these cases there is no type-shift. The nouns, verbs and adjectives all refer to situations and events or 2ndOrderEntities. Type shifting relations across part of speech, such as between the *cutting event* and the *cutting instrument,* will be discussed as ROLEs.

Type-persistent relations across parts-of-speech can be tested using frames that explicitly compensate for the syntactic differences between the word pairs. The following tests express a synonymy relation between nouns and verbs in general:

Test 4		**XPOS_Near_Synonymy between nouns and verbs**		
yes	a	If there is a case of a/an X then something/someone/it Ys		
yes	b	If something/someone/it Ys then there is a case of a/an X		
Conditions:		- X is a noun in the singular		
		- Y is a verb in the third person singular form		
		- there are no specifying PPs that apply to the X-phrase or the Y-phrase		
		- preferably there is a morphological link between the noun and the verb		
Example:	a	If there is a case of a movement then something moves		
	b	If something moves then there is a case of a movement		
Effect:		movement N	XPOS_NEAR_SYNONYM	move V
		move V	XPOS_NEAR_SYNONYM	movement N

The distinction between hyponymy and synonymy is not always clear-cut. Sometimes concepts can be very close showing only a very limited specialization. In the case of relations across part-of-speech we can at least formulate the extra conditions that a strong morphological link between the two words is preferred, as is here the case for *movement* and *move*.

Whereas the previous test works both for non-dynamic states and dynamic events, the next tests only apply to dynamic or static events:

Test 5		**XPOS_Near_Synonymy between event-denoting nouns and verbs**		
yes	a	if a(n) X takes place then something/somebody/it Ys		
yes	b	if something/somebody/it Ys then a/an X takes place		
Conditions:		- X is a noun in the singular		
		- Y is a verb in the third person singular form		
		- there are no specifying PPs that apply to the X-phrase or the Y-phrase		
		- preferably there is a morphological link between the noun and the verb		
Example:		X = movement		
		Y = move		
Effect:		movement N	XPOS_NEAR_SYNONYM	move V
		move V	XPOS_NEAR_SYNONYM	movement N

Test 6		**XPOS_Near_Synonymy between state-denoting nouns and verbs**		
yes	a	If there is a state of X then something/someone/it Ys		
yes	b	If something/something/it Ys then there is a state of a/an X		
Conditions:		- X is a noun in the singular		
		- Y is a verb in the third person singular form		
		- there are no specifying PPs that apply to the X-phrase or the Y-phrase		
		- preferably there is a morphological link between the noun and the verb		
Example:				
yes	a	If there is a state of sleep then something/someone/it sleeps		
yes	b	If something/something/it sleeps then there is a state of a/an sleep		
Effect:		sleep N	XPOS_NEAR_SYNONYM	sleep V
		sleepV	XPOS_NEAR_SYNONYM	sleep N
Example:	a	If something/someone/it exists then there is a state of existence		
	b	If there is a state of existence then something/someone/it exists		
Effect:		to exist (X)	XPOS_NEAR_SYNONYM	existence (Y)
		existence (Y)	XPOS_NEAR_SYNONYM	to exist (X)

The next tests are similar to the previous ones but apply to adjectives/adverbs and nouns or verbs that denote non-dynamic states. The test is only different in so far that adjectives/adverbs need a copula to occur in the same sentence as above:

Test 7		**XPOS_Near_Synonymy between state-denoting nouns and adjectives/adverbs**
yes	a	If there is a state of X then something/someone/it is Y
yes	b	If something/someone/it is Y then there is a state of a/an X
Conditions:		- X is a noun in the singular - Y is an adjective - there are no specifying PPs that apply to the X-phrase or the Y-phrase - preferably there is a morphological link between the noun and the adjective
Example:	a	If there is a state of popverpty then something/someone/it is poor
	b	If something/someone/it is poor then there is a state of a/an poverty
Effect:		poverty N XPOS_NEAR_SYNONYM poor A poor A XPOS_NEAR_SYNONYM poverty N

Test 8		**XPOS_Near_Synonymy between state-denoting verbs and adjectives/ adverbs**
Yes	a	if something/someone/it Xs then something/someone/it is Y
Yes	b	if something/someone/it is Y then something/someone/it Xs
Conditions:		- X is a verb in the third person singular form - Y is an adjective - there are no specifying PPs that apply to the X-phrase or the Y-phrase - preferably there is a morphological link between the noun and the verb
Example:	a	if someone/it lives then someone/it is alive
	b	if someone/it is alive then someone/it lives
Effect:		to live (X) XPOS_NEAR_SYNONYM alive (Y) alive (Y) XPOS_NEAR_SYNONYM to live (X)

2.2.3.2. Hyponymy

As argued in Fellbaum (1998), hyponymy is the most fundamental relation around which the wordnets are constructed. Chains of hyponymy relations such as:

taxi HAS_HYPERONYM car HAS_HYPERONYM motor vehicle HAS_HYPERONYM vehicle HAS_HYPERONYM instrument HAS_HYPERONYM object HAS_HYPERONYM entity

can form the backbone of a knowledge base or lexicon, via which rich semantic specifications can be inherited in a consistent way to thousands of more specific concepts. In WordNet, multiple hyperonyms have occasionally been encoded. In EuroWordNet, we have tried to encode multiple hyponymy relations more comprehensively. However, hierarchical structures quickly become very complex once this is allowed, and consistency should be checked by actually implementing and applying inheritance. Any hierarchical structure should therefore be populated with features that can be tested against a corpus or by some task, to verify its quality.

Hyperonymy and hyponymy are inverse relations, which roughly correspond to the notion of class-inclusion: if Y is a kind of X, then X is hyperonym of Y and Y is an hyponym of X. Both relations are asymmetric and transitive. A hyponymy relation implies that the hyperonym (the more general class) may substitute the hyponym (the more specific subtype) in a referential context but not the other way around. A referential context is a context where only the denotational range (the set of discourse entities) is considered (grammatical, register, pragmatic and other non-semantic properties of the considered words or context are neglected). Given these constraints there must be a full inclusion of the set of entities denoted by the hyponym in the set of entities denoted by the hyperonym. An extra constraint can be that there must be multiple co-hyponyms to result in a genuine hyponymy relation. This means that the denotation of the hyponym is never equal to the denotation of the hyperonym, i.e. it must be a proper subset.

The same substitution principle as discussed above for synonymy can thus be applied to hyponymy relations but it only holds in one direction. However, to more clearly elicit the difference in specificity the tests have been extended with general specifying phrases. In addition to the formal substitution-sentences we can state that:

> If a pair of words W1 and W2 fits the test frame then there should be at least one other word W3 which fits this frame in relation to W2 so that W1 and W3 are so-called co-hyponyms of W2. The presence of co-hyponyms is a necessity to establish a genuine hyponymy relation.

In the next test three different paraphrases are used which elicit the implicational relation between the hyponym and the hyperonym.

Test 9		**Hyponymy-relation between nouns**		
yes	a	A/an X is a/an Y with certain properties)		
		It is a X and therefore also a Y		
		If it is a X then it must be a Y		
no	b	the converse of any of the (a) sentences.		
Conditions:		- both X and Y are singular nouns or plural nouns		
Example:	a	A car is a vehicle with certain properties		
	b	?A vehicle is a car with certain properties		
	a	It is a car and therefore also a vehicle		
	b	?It is a vehicle and therefore also a car		
	a	If it is a car then it must be a vehicle		
	b	?If it is a vehicle then it must be a car		
Effect:		car N	HAS_HYPERONYM	vehicle N
		vehicle N	HAS_HYPONYM	car N

Without the specifying phrase, this test can also be used for synonymy. The next test indicates a more specific type of hyponymy between kinds, species, races and brands:

Test 10		**Hyponymy-relation between nouns of species and classes, which is reflected by the explicit hyponymy nouns such as sort/kind, type, race, species.**		
yes	a	A/an X is a kind/type/race/species/brand of Y(s)		
no	b	the converse of the (a) sentence.		
Conditions:		- X is a singular noun		
		- Y is a singular or plural noun		
Example:	a	A mercedes is a kind of car		
	b	?A car is a kind of mercedes		
Effect:		mercedes N	HAS_HYPERONYM	car N
		car N	HAS_HYPONYM	merdeces N

This test cannot be used for synonymy.

A general criterion for testing hyponymy between verbs is the following:

> A verb synset X is a hyponym of another verb synset Y (and, by the same token, Y a hyperonym of X) if *He is X-ing* entails but is not entailed by *He is Y-ing*.

The following sentences then should be true and false respectively:

> - He Vs_1 therefore he Vs_2 yes
> - He Vs_2 therefore he Vs_1 no
> Clear yes-no = V_1 is a hyponym of V_2 (and V_2 is a hyperonym of V_1)

This general test is however not sufficient, because it does not distinguish between verbs connected by a hyponymy relation and verbs connected by a more general entailment relation. In fact, in this test, V_1 could be, for instance, *to snore* and V_2 could be *to sleep* (indeed, *He is snoring* entails but is not entailed by *He is sleeping*), which are not connected by a hyponymy relation. The test should be reformulated as a more specific phrase. Since each hyponym is equivalent to a paraphrase in which its hyperonym is syntagmatically modified, we can state the following formal criteria for the definition of hyperonymy/hyponymy:

Test 11		**Hyperonymy/hyponymy between verb synsets**		
yes	a	to X is to Y + AdvP/AdjP/AdjP/NP/PP		
no	b	to Y is to X + AdvP/AdjP/NP/PP		
Conditions:		- X is a verb in the infinitive form		
		- Y is a verb in the infinitive form		
		- there is at least one specifying AdvP, NP or PP that applies to the Y-phrase		
Example:	a	to run is to go fast		
	b	* to go is to run fast		
Effect:		{to run} (X)	HAS_HYPERONYM	{to go} (Y)
		{to go} (Y)	HAS_HYPONYM	{to run} (X)

As is the case for near_synonymy, hyponymy can also be established between words with different parts of speech. This relation also come in inverse pairs. In the previous section, we have seen some test sentences for synonymy-relations across parts-of-speech. In principle these tests can also be used as a basis for hyponymy-tests with some additions to elicit the difference in specificity:

Test 12		**XPOS_Hyponymy of nouns and verbs denoting events**		
yes/no	a	If a/an X takes place then something/someone/it Ys + NP, PP (in a certain way)		
no/yes	b	If something/someone/it Ys then there is a/an X takes place		
Conditions:		- X is a noun in the singular		
		- Y is a verb in the third person singular form		
		- there should be at least one specifying NP or PP that makes the Y-phrase equivalent to the X-phrase or the other way around.		
		- preferably there is no morphological link between the noun and the verb		
Example:	a	If an election takes place, then somebody votes for a political party		
no/yes	b	If someone votes for a political party then an election takes place		
Effect:		{election}N (X)	HAS_XPOS_HYPERONYM	{to vote}V (Y)
		{to vote} V (Y)	HAS_XPOS_HYPONYM	{election}(X)

The reversal of the score leads to a reversion of the hyponymy: noun-to-hyperonym-verb or verb-to-hyperonym-noun. As long as one direction has a clear positive score and the other direction has a clear negative score we are dealing with a hyponymy relation.

The next test only applies nouns and verbs expressing non-dynamic situations or states:

Test 13		**XPOS_Hyponymy between state-denoting nouns and verbs**		
yes/no	a	if there is a state of X then something/someone Ys + NP, PP (in a certain way)		
no/yes	b	if someone/something/it Ys then a state of a/an certain X applies		
Conditions:		- X is a noun in the singular		
		- Y is a verb in the third person singular form		
		- there should be at least one specifying NP or PP that makes the Y-phrase equivalent to the X-phrase or the other way around.		
		- preferably there is no morphological link between the noun and the verb		
Example:	a	If there is a state of paranoia then someone fears something intensively		
	b	* If someone fears something then there is a certain state of paranoia		
Effect:		paranoia (Y)	HAS_XPOS_HYPERONYM	to fear (X)
		to fear (X)	HAS_XPOS_HYPONYM	paranoia(Y)

The next test elicits hyponymy between adjectives/adverbs and nouns that denote non-dynamic situations or states:

Test 14		**XPOS_Hyponymy between state-denoting nouns and adjectives**
yes/no	a	if there is a state of X then something/someone/it is Y in a certain way
no/yes	b	if something/someone/it is Y then a state of a/an X applies
Conditions:		- X is a noun in the singular - Y is an adjective - there is at least one specifying adverb, NP or PP that applies to the X-phrase or the Y-phrase - preferably there is a no morphological link between the noun and the adjective
Example:	a	If there is a state of brain-death then someone is dead in a certain way
	b	*if something/someone/it is dead then a state of a/an brain-death applies
Effect:		brain-death (Y) HAS_XPOS_HYPERONYM dead (X) dead (X) HAS_XPOS_HYPONYM brain-death(Y)

Note that the XPOS_HYPONYMY relation can also be used to relate nouns that head a class of adjectival values:

size XPOS_HYPONYMs small, big, medium.
colour XPOS_HYPONYMs black, white, blue, green, yellow, red.
taste XPOS_HYPONYMs sour, sweet, bitter.
shape XPOS_HYPONYMs round, rectangular, cubic, triangular, oval.

In WordNet1.5, these cases are related by the ATTRIBUTE relation between nouns and adjectives.

Finally, the next test elicits hyponymy between static verbs and adjectives:

Test 15		**Xpos_Hyponymy between state-denoting verbs and adjectives/adverbs**
yes	a	If something/someone/it is Y then something/someone/it Xs + AdvP/AdjP/NP/PP
no	b	If something/someone/it Xs then something/someone/it is in a certain state of being Y
Conditions:		- X is a verb in the third person singular form - Y is an adjective - there is at least one specifying AdvP, NP or PP that applies to the X-phrase - preferably there is no morphological link between the adjective and the verb
Example:	a	If someone is horrified then someone fears something intensively
	b	* If someone fears something then someone is in a certain state of being horrified
Effect:		horrified (Y) HAS_XPOS_HYPERONYM to fear (X) to fear (X) HAS_XPOS_HYPONYM horrified (Y)

2.2.3.3. Antonymy

Antonymy relates lexical opposites, such as “to ascend” and “to descend”, “good” and “bad” or “justice” and “injustice”. It is clear that antonymy is a symmetric relation, but little more can be said, since it seems to encode a large range of phenomena of opposition, e.g. “rich” and “poor” are scalar opposites with many values in between the extremes, “dead” and “alive” can be seen as complementary opposites (Cruse 1987).

It is also unclear whether antonymy stands between either word forms or word meanings. For instance, “appearance” and “arrival” are, in the appropriate senses, synonyms; but linguistic intuition says that the appropriate antonyms are different for each word (“disappearance” and “departure”). With respect to this, EWN will assume the solution adopted by Miller's WordNet, that is, antonymy is considered to be a relation between word forms, but not between word meanings -namely synsets. Therefore, in the example above, the antonymy relation will hold between “appearance” and “disappearance”, “arrival” and “departure” as word forms. In those cases that antonymy also holds for the other variants of the synset we use a separate NEAR_ANTONYM relation. Finally, we may find cases in which there is an opposition between synsets with different parts-of-speech. Just as with the synonymy and hyponymy relations we store these relations as XPOS_NEAR_ANTONYM relations.

Antonyms typically form contrasting categories within the same dimension. This means that an Antonym not only contrasts with another antonym in one or more features (e.g. animate/inanimate) but that they have to share the same hyperonym: i.e. they have to be competitors within a reasonable denotational range. This latter criterion prevents us from contrasting irrelevant pairs such as “car” and “love”. An antonymy test therefore has to consist of two parts: one part expressing the contrast and one part expressing the shared dimension or hyperonym

Test 16		**Antonymy between nouns**		
yes	a	X and Y are both a kind of Z but X is the opposite of Y		
yes	b	the converse of (a)		
Conditions:		- X and Y are singular or plural nouns		
		- Z is a hyperonym of both X and Z and within a reasonable, competitive denotational range.		
Example:	a	man and woman are both a kind of human being but man is the opposite of woman		
	b	woman and man are both a kind of human being but woman is the opposite of man		
Effect:		man- N	ANTONYM	woman-N
		woman-N	ANTONYM	man-N

Verbal opposition is often revealed by morphological structure: *tie/untie*, *appear/disappear*, *approve/disapprove*, etc. However, in other cases, the antonymy rises from the opposition between adjectives or direction incorporated within the meaning of verbs, e.g. in Italian: *abbellire/imbruttire* (*prettify/uglify*), *dimagrire/ingrassare* (*slim/fat*), *entrare/uscire* - to go in/to go out, *salire/scendere* - to go up/to go down). Finally, a special class of verbal antonyms in WN 1.5 occur within the same semantic field and “refer to the same activity, but from the viewpoint of different participants” (Fellbaum 1990:51): *lend/borrow*, *teach/learn*, *buy/sell*, etc.

Test 17		**Antonymy between verb**		
yes	a	If something/someone/it Xs then something/someone/it does not Y		
yes	b	If something/someone/it Ys then something/someone/it does not X		
Conditions:		- X is a synset variant in the third person singular form		
		- Y is a synset variant in the third person singular form		
	i.	- X and Y are members of co-hyponym synsets		
	ii.	- there is a hyperonym of X which is opposite to a hyperonym of Y		
	iii.	- the situation referred to by X has an addressee and the addressee is the protagonist of the situation referred to by Y		
Example:	ia	If he gets fat then he does not get thin		
	ib	If he gets thin then he does not get fat		
	iia	If he sells then he does not buy		
	iib	If he buys then he does not sell		
	iiia	If he gives then he does not take		
	iiib	If he takes then he does not give		
Effect:		{to get fat, to put on weight}		
			NEAR_ANTONYM	{to get thin, to lose weight}
		{to sell, to exchange for money}		
			NEAR_ANTONYM	{to buy, to purchase, to take}
		{to give}		
			NEAR_ANTONYM	{to take, to take away}

If the antonymy relation holds between all variants, the relation is NEAR_ANTONYM, otherwise it is ANTONYMY. Antonymy between different POS is only allowed between synsets (and not variants):

Test 18		**XPOS_Antonymy between dynamic verbs and nouns**
yes	a	If something/someone/it Xs then a/an Y does not take place
yes	b	If a/an Y takes place then something/someone/it does not X
Conditions:		- X is a verb in the third person singular form - Y is a noun in the singular - X and Y are (XPOS) co-hyponyms
Example:	a	If someone falls asleep then awakening does not take place
	b	If awakening takes place then someone does not fall asleep
Effect:		{to fall asleep} (X) XPOS_NEAR_ANTONYM {awakening} (Y)

Test 19		**XPOS_Antonymy between static verbs and nouns**
yes	a	If something/someone/it Xs then something/someone/it is not in a state of Y
yes	b	If something/someone/it is in a state of Y then something/someone/it does not X
Conditions:		- X is a verb in the third person singular form - Y is a noun in the singular - X and Y are (XPOS) co-hyponyms
Example:	a	If someone loves someone then someone is not in a state of hate
	b	If someone is in a state of hate then someone is not loving
Effect:		{to love} (X) XPOS_NEAR_ANTONYM {hate} (Y)

Test 20		**Antonymy between verbs and adjectives (or adverbs)**
yes	a	If something/someone/it Xs then something/someone/it is not Y
yes	b	If something/someone/it is Y then something/someone/it does not X
Conditions:		- X is a verb in the third person singular form - Y is an adjective - X and Y are (XPOS) co-hyponyms
Example:	a	If someone sleeps then someone is not awake
	b	If someone is awake then someone does not sleep
Effect:		{to sleep} (X) XPOS_NEAR_ANTONYM {awake} (Y)

2.2.3.4. Meronymy

Most scholars in Lexical Semantics (e.g. Cruse, 1986) and Psycholinguistics (e.g. Winston et al. 1987) also claim that the so-called Part-Whole relation is a family of relations. The most salient subtypes are:

(i) between (the nouns standing for) a whole and their constituent parts (“part”, e.g. “hand”-“finger”);
(ii) between a portion and the whole from which it has been detached (“portion”, e.g. “ingot”-“metal”);
(iii) between a place and a wider place which includes it (“location”, e.g. “oasis”-“desert”);
(iv) between a set and their members (e.g. “fleet” -“ship”);
(v) between a thing and the substance it is made of (“made-of”, e.g. “book”-“paper”).

In EuroWordNet, we decided to limit part-whole relations to these five types. A general unspecified relation is used to cover unclear cases. A further differentiation is made between unique and non-unique parts. Unique parts belong to one type of whole, e.g. *finger* which is only a part of *hand*, non-unique parts can belong to a diverse range of wholes, e.g. *window* which can be a part of a *building*, *vehicle*, *container*, etc.. Whether or not a part is unique follows from the fact that there are multiple disjunctive wholes to which it is linked.

Also the Part-Whole relations come in inverse pairs, namely holonym and meronym - if X is the holonym of Y, Y is the meronym of X. Likewise, we defined one general relation HAS_HOLONYM (and its inverse HAS_MERONYM) and five subtypes of them, namely

- HAS_HOLO_PART and HAS_MERO_PART
- HAS_HOLO_PORTION and HAS_MERO_PORTION
- HAS_HOLO_LOCATION and HAS_MERO_LOCATION
- HAS_HOLO_MEMBER and HAS_MERO_MEMBER
- HAS_HOLO_MADEOF and HAS_MERO_MADEOF

As explained above, the automatically reversed relations will get the label **reversed**. In the examples below we will not express this because the tests do not make clear which direction of the relation was explicitly coded and which direction was the result of automatic reversal.

Test 21		**General meronymy for nouns**		
yes	a	(a/an) X makes up a part of (a/an) Y		
		(a/an)Y has (a/an) Xs		
no	b	the converse of the a) relations		
Conditions:		X and Y are concrete nouns and are interpreted generically		
Effect:		X	HAS_HOLONYM	Y
		Y	HAS_MERONYM	X

Test 22		**MEMBER/GROUP meronymy for nouns using a relational member-noun**		
yes	a	(a/an) X is a member/element of (a/an/the) Y		
no	b	the converse of a)		
Conditions:		- X is a single object-denoting noun		
		- Y is a multiform noun (either a group-noun, a collective-noun or as a lexicalized plural denoting multiple objects)		
		- preferably humans, animals, plants or vehicles or closed sets such as the number system, or the alphabet.		
Example:	a	a player is a member of a team		
	*b	a team is a member of a player		
Effect:		player	HAS_HOLO_MEMBER	team
		team	HAS_MERO_MEMBER	player

Several studies suggested that the *portion-of* relation differs in several aspects from other meronymy relations:

(i) the whole always pre-exist the portion;
(ii) usually portions (as concepts) do not receive a separate lexical item but are realized by sense extension (for instance, there is no lexical item equivalent to “portion of cake”);
(iii) boundaries of portions usually are not defined;

Sometimes portions are sufficiently common in a particular language to become lexicalized. These lexical items will be linked to their wholes by means of a *has_holo_portion* link according to the following test:

Test 23		**PORTION meronymy for nouns using a relational amount-noun**		
yes	a	(a/an) X is an (amount/piece/portion) of Y		
no	b	the converse of (a)		
Conditions:		X and Y are substance denoting nouns		
Example:	a	a drop is an amount of liquid		
	*b	a liquid is an amount of a drop		
Effect:		drop	HAS_MERO_PORTION	liquid
		liquid	HAS_HOLO_PORTION	drop

The *has_holo/mero_part* relation typically relates components to their wholes, namely: something which is either topologically or temporally included in a larger entity and which as well bears some kind of autonomy (non-arbitrary boundaries) and a definite function with respect to the whole.

Test 24		**PART meronymy for nouns**		
yes	a	a/an X is a component of a/an Y		
yes	b	a/an Y is a whole/system/complex/network/arrangement/construction of parts/components among which a/an X		
Conditions:		X and Y are concrete nouns denoting objects, there must be several Xs		
Example:	a	a wheel is a component of a car		
	*b	a car is a component of a wheel		
Effect:		wheel	HAS_HOLO_PART	car
		car	HAS_MERO_PART	wheel

The condition states that there must be multiple components (which can be of the same type) and that both the holonym and the meronym should be concrete objects. Complex holonyms can also contain substances but in that case the MADE_OF relation is used.

There are two basic ways of viewing entities in the world, namely either as an individuated thing or as the stuff from which they are made of. This way, for instance a book can be alternatively named “a book” or “paper”. The relation between things and the stuff which compose them is called MADE_OF. It is defined by the suitability of the following test:

Test 25		**MADEOF meronymy for nouns**		
yes	a	a/an X is made of Y		
no	b	the converse of (a)		
Conditions:		- X is a concrete object		
		- Y is a concrete substance		
Example:	a	a stick is made of wood		
	*b	wood is made of stick		
Effect:		stick	HAS_MERO_MADEOF	wood
		wood	HAS_HOLO_MADEOF	stick

Place nouns form an important set in a lexical database. Space, in a general sense, is by definition contiguous and the sub-division in more inclusive pieces of space largely seems to be a matter of lexicalisation. Nouns for places must stand in a relation of lexical-semantic inclusion to the nouns of the larger places which include them; a relation which is parallel to the topological 'real-world' relation which stands between the places named.

Test 26		**LOCATION meronymy for nouns**		
yes	a	(a/an/the) X is a place located in (a/an/the) Y		
no	b	the converse of (a)		
Conditions:		- X is a concrete noun		
		- Y is a concrete noun		
Example:	a	the centre is a place located in a city		
	*b	the city is a place located in a centre		
Effect:		centre	HAS_HOLO_LOCATION	city
		city	HAS_MERO_LOCATION	centre

2.2.3.5. ROLE and INVOLVED

So far, all relations that have been discussed are between entities of the same paradigmatic type. Synonymy, hyponymy, antonymy and meronymy (within or across part-of-speech) can only be expressed either between pairs of 1st, 2nd or pairs of 3rdOrderEntities respectively, but never across these types. All these relations are therefore type-persistent. In this section we will describe the relations that can only be expressed across different ontological types, more specifically, the different roles and functions that 1st and 3rdOrderEntities may have in events (2ndOrderEntities).

From a cognitive point of view, function is one of the major features that organizes human knowledge. Likewise, functionality is widely reflected in the lexicon. Languages are rich in derivational procedures that generate nouns from verbs or the other way round along a functional dimension -e.g. *run/runner*, *telephone/to telephone*. In such cases, there is a tight semantic relation between both lexical units that is potentially useful for linguistic engineering tasks. Functional relations are often related to telicity but, since they also cover other aspects of semantic entailment, they will be referred to as - more generically - *involvement* relations.

If the relation goes from a concrete or mental entity (only nouns denoting 1st or 3rdOrderEntities) to verbs or event denoting nouns (2ndOrderEntities), it will be called *role*, the inverse from events (2ndOrderEntities) to concrete or mental entities (nouns) is called *involved*. For instance, the verb *to hammer* will directly be linked to the noun *hammer* by means of the INVOLVED_INSTRUMENT relation and the latter will be related back by a ROLE_INSTRUMENT relation to the verb. Similarly, the noun *carpenter* can be connected with the verb *to hammer* by means of the ROLE_AGENT relation, and the correspondent link from the verb to the noun (i.e., *to hammer* --> INVOLVED_AGENT --> *carpenter*) is then automatically derived. The verb hammer will thus have several INVOLVED relations, some of them being labelled as reversed, others perhaps as disjuncts (e.g. multiple *agents* connected to it).

Although the ROLE/INVOLVED relations often correlate with the kind of arguments that a verb requires as its complements, they do not necessarily coincide with them. For instance, a verb like *to move*, in its inchoative sense, allows both agent and patient arguments, but has no particular ‘involved-agent’ or ‘involved-patient’ in its meaning. That is, the meaning of the verb does not motivate a link to any specific involved-argument. On the other hand, a verb like *sgambettare* (an Italian verb meaning *to kick one’s legs about* and only used to refer to a movement performed by babies) does incorporate a specific ‘agent-protagonist’ which differentiates it from other movements. This will be encoded by means of the relation INVOLVED_AGENT --> *babies*

Also note that ROLE/INVOLVED relations are not the same as selectional restrictions. The instrument of “to hammer” can be any physical object and is not necessarily restricted to the instrument “hammer”. However, the relation to the instrument “hammer” is a conceptually salient and will immediately be triggered regardless of the context.

In addition to the general relation ROLE/INVOLVED, we distinguished: AGENT, PATIENT, INSTRUMENT, RESULT, LOCATION, DIRECTION, SOURCE_DIRECTION, TARGET_DIRECTION, where each relation is differentiated in both direction as a ROLE and an INVOLVEMENT. The differentiation is based on the need for these relations to encode and clarify concepts in the processed lexicons. There is no fundamental reason for making this choice or for not distinguishing more relations.

Just as with the meronymy relations, the general relation ROLE/INVOLVED is used for cases where the tests or the criteria for extracting these relations from resources cannot discriminate between the subtypes. The general test for a ROLE/INVOLVED relation is as follows:

Test 29	**INVOLVED/ROLE as general relation**		
yes	(a/an) X is the one/that who/which is typically involved in Ying		
Conditions:	X is a noun		
	Y is a verb in the infinitive form		
Example:	A hammer is that which is typically involved in hammering		
Effect:	{hammer} (X)	ROLE	{to hammer} (Y)
	{to hammer} (Y)	INVOLVED	{hammer} (X)

The next tests can then be used to elicit more specific involvements.

The first two relations AGENT and PATIENT are based on the notions of 'proto-agent' and 'proto-patient' as defined by Dowty (1988). According to Dowty, various properties implied within the meaning of a verb contribute to the definition of proto-roles:

(1) Typical properties for the Agent Proto-Role:
- a. volition
- b. sentience (and/or perception)
- c. causes event
- d. movement

(2) Typical properties for the Patient Proto-Role:
- a. change of state (including come-into-being, cease-to-be)
- b. incremental theme
- c. causally affected by event
- d. stationary

A proto-agent does not need to have all the properties indicated, but, among the arguments of a verb, it is the one which has more proto-agent properties. The following tests can be used to elicit typical agents and patients in general:

Test 28		**Agent Involvement**		
yes	a	(A/an) X is the one/that who/which does the Y, typically intentionally.		
Conditions:		- X is a noun - Y is a verb in the gerundive form		
Example:	a	A teacher is the one who does the teaching intentionally		
Effect:		{to teach} (Y)	INVOLVED_AGENT	{teacher} (X)

Test 29		**Patient Involvement**		
yes	a	(A/an) X is the one/that who/which undergoes the Y		
Conditions:		- X is a noun - Y is a verb in the gerundive form		
Example:	a	A learner is the one who undergoes the learning		
Effect:		{to learn} (Y)	INVOLVED_PATIENT	{learner} (X)

RESULTs are a special kind of PATIENTs. In this case, the entity is not jut changed or affected but it comes into existence as a result of the event:

Test 30		**Result Involvement**		
yes	a	(A/an) X is comes into existence as a result of Y		
yes	b	(A/an) X is the result of Y		
yes	c	(A/an) X is created by Y		
Conditions:		- X is a noun - Y is a verb in the gerundive form and a hyponym of “make”, “produce”, “generate”.		
Example:	a	a crystal comes into existence as a result of crystalizing		
	b	a crystal is the result of crystalizing		
	c	a crystal is created by crystalizing		
Effect:		{to crystalize} (Y)	INVOLVED_RESULT	{crystal} (X)

Note that RESULTs are strictly concrete entities (1stOrder) or mental objects such as ideas (3rdOrder). Situations that *result* from other situations are related by the CAUSE relation (see below). Furthermore, the event should be a resultative verb, i.e. a hyponym of concepts such as *make, produce, generate*.

A different type of relation is INSTRUMENT, which mostly applies to inanimate entities used by animate entities to get some effect or result:

Test 31		**Instrument Involvement**		
yes	a	(A/an) X is either i) the instrument that or ii) what is used to Y (with)		
Conditions:		- X is a noun - Y is a verb in the infinitive form		
Example (1):		An hammer is the instrument that is used to hammer		
Effect:		{hammer} (X)	ROLE_INSTRUMENT	{to hammer} (V)
Effect:		{to hammer} (Y)	INVOLVED_INSTRUMENT	{hammer} (X)
Example (2):		A sailing boat is what is used to sail with		
Effect:		{sail} (X)	ROLE_INSTRUMENT	{to ail} (V)
Example (1):		Pen/Ink/Paper is what is used to write		
Effect:		{pen} (X)	ROLE_INSTRUMENT	{to write} (X)
		{ink} (X)	ROLE_INSTRUMENT	{to write} (X)
		{paper} (X)	ROLE_INSTRUMENT	{to write} (X)

Two types of location involvements are distinguished. The place where something takes place is called LOCATION and the place to or from where movement is directed is called DIRECTION:

Test 32		**Location Involvement**		
yes	a	(A/an) X is the place where the Y happens		
Conditions:		- X is a noun		
		- Y is a verb in the gerundive form		
Example:	a	A school is the place where the teaching happens		
Effect:		{school} (X)	ROLE_LOCATION	{to teach} (Y)
		{to teach} (Y)	INVOLVED_LOCATION	{school} (X)

Test 33		**Direction Involvement**		
yes	a	It is possible to Y from/to/over/across/through a place (X)		
Conditions:		- Y is a verb in the infinitive form		
Example:	a	It is possible to pass though a place		
Effect:		{to pass} (Y)	INVOLVED_DIRECTION	{place} (X)

The DIRECTION relation is then further differentiated into:

Test 34		**Source-Direction Involvement**		
yes	a	(A/an/the) X is the place from where Ying begins/starts/happens / one Ys		
Conditions:		- X is a noun		
		- Y is a verb		
Example:	a	The start is the place from where the racing starts		
Effect:		{to race} (Y)	INVOLVED_SOURCE	{the start} (X)

Test 35		**Target-Direction Involvement**
yes	a	(a/an/the) X is the place to which Ying happens / one Ys
Conditions:		- X is a noun
		- Y is a verb
Example:	a	The ground is the place to which one collapses/falls heavily
Effect:		{to collapse, to fall heavily} (Y)
		INVOLVED_TARGET_DIRECTION
		{ground} (X)

The INVOLVED_DIRECTION relation is useful to distinguish different incorporations in a language (e.g., the Italian verb *nuotare* (to swim) has no INVOLVED_DIRECTION) and among differences of lexicalisation across languages (e.g., *to swim* has a generic INVOLVED_DIRECTION).

2.2.3.6. CO_ROLE

Especially in Germanic languages, many compounds are lexicalized that incorporate different participants of an event in their meaning, but the event itself is not made explicit, e.g.: *guitar player* or *ice saw*. In some cases the event is lexicalized as a specific verb but still often only one of the components is related to the verb, i.e. a *saw* as an instrument of *to saw* but *ice* is not a typical patient of *saw*. The concept *ice* is only related to *saw* via *ice-saw*, there is no other reason to link *ice* and *saw*. To properly relate these compounds we would thus directly want to link the co-participants. This can be done using the so-called CO_ROLE relation. CO_ROLES represent pairs of ROLE relations between concrete and/or mental entities, while the event itself is not necessarily made explicit (although it may be).[7] CO_ROLES are thus partially type-persistent: there may be co_roles between 1st and 3rdOrderEntities (e.g. thinker CO_AGENT_RESULT thought) but not between 1st and 2nd or 3rd and 2ndOrderEntities. Given the above ROLE relations we thus get the following CO_ROLEs:

CO_ROLE (general relation that is bi-directional)
CO_AGENT_PATIENT & CO_PATIENT_AGENT
CO_AGENT_INSTRUMENT & CO_INSTRUMENT_AGENT
CO_AGENT_RESULT & CO_RESULT_AGENT
CO_PATIENT_INSTRUMENT & CO_INSTRUMENT_PATIENT
CO_PATIENT_RESULT & CO_RESULT_PATIENT
CO_INSTRUMENT_RESULT & CO_RESULT_INSTRUMENT

[7] An alternative would be to use 3-place relations: ice-saw ROLE_INTRUMENT saw INVOLVED_PATIENT ice. These are however not foreseen in the database.

Note that there is no corresponding CO_ROLE relation to ROLE_LOCATION and ROLE_DIRECTION. The reason for this is that the relation would overlap too much with HAS_HOLO_LOCATION. If some entities is involved in an event at some location, then this entity can also be located at that location during the event, and hence the HAS_HOLO_LOCATION relation holds between this entity and the location.

The above examples will then be encoded as follows:

guitar player
    HAS_HYPERONYM player
    CO_AGENT_INSTRUMENT guitar
player
    HAS_HYPERONYM person
    ROLE_AGENT to play music
    CO_AGENT_INSTRUMENT musical instrument
to play music
    HAS_HYPERONYM to make
    ROLE_INSTRUMENT musical instrument
guitar
    HAS_HYPERONYM musical instrument
    CO_INSTRUMENT_AGENT guitar player
ice saw
    HAS_HYPERONYM saw
    CO_INSTRUMENT_PATIENT ice
saw
    HAS_HYPERONYM saw
    ROLE_INSTRUMENT to saw
ice
    CO_PATIENT_INSTRUMENT ice saw REVERSED

Examples of the other relations are:

criminal
    CO_AGENT_PATIENT victim
novel writer/ poet
    CO_AGENT_RESULT novel/ poem
pastry dough/ bread dough
    CO_PATIENT_RESULT pastry/ bread
photograpic camera
    CO_INSTRUMENT_RESULT photo

We will not give specific tests for the CO_ROLE relations. The above ROLE/INVOLVED test can be used in combination to verify a CO_ROLE relation.

2.2.3.7. CAUSES and IS_CAUSED_BY

The causal relation is used in WN1.5 for verb pairs such as *show/see*, *fell/fall*, *give/have*. Fellbaum (1990: 54) states that the causal relation only holds between verbs, and only between verbs that are temporally disjoint. In EuroWordNet, the cause relation is used to link 2ndOrderEntities, which can be either verbs, nouns and adjectives (the relation is thus type-persistent but can apply across POSs). The only constraint is that the causing event should be dynamic (henceforth 'dynamic situations' or *dS*), whereas the resulting situation can either be static or dynamic. In addition, we distinguish among 3 temporal relationships between the (dynamic/non-dynamic) situations related by cause:

- a cause relation between two situations which are temporally disjoint: there is no time point when dS_1 takes place and also S_2 (which is caused by dS_1) and vice versa (e.g., in the case of *to shoot/to hit*);
- a cause relation between two situations which are temporally overlapping: there is at least one time point when both dS_1 and S_2 take place, and there is at least one time point when dS_1 takes

place and S_2 (which is caused by dS_1) does not yet take place (e.g., in the case of *to teach/to learn*);
- a cause relation between two situations which are temporally co-extensive: whenever dS_1 takes place also S_2 (which is caused by dS_1) takes place and there is no time point when dS_1 takes place and S_2 does not take place, and vice versa (e.g., in the case of *to feed/to eat*).

If situations are co-extensive it may be argued that we are not dealing with two separate events at all, e.g. "to dig" and "to dig a hole". In that case, we may also be dealing with a hyponymy relation between one verb which is simply more inclusive (implying a result) than another verb (change without necessarily implying a result). We decided to prefer hyponymy above cause when non-disjoint verb-pairs also pass the hyponymy test.

As we have already recalled, then, different types of causality can also be distinguished with respect to the factivity of the effect. In the following general formal criteria for the definition of causation relation are provided.

Test 36		**Factive causation relation**			
yes	a	(To/A/an) X causes (to/a/an) Y to take place			
		(To/A/an) X has (to/a/an) Y as a consequence			
		(To/A/an) X leads to (to/a/an) Y			
no	b	the converse of (a)			
Conditions:		- X is a verb in the infinitive form or X is a noun in the singular			
		- Y is a verb in the infinitive form or Y is a noun in the singular			
Example:	a	to kill (/a murder) causes to die (/ death)			
		to kill (/a murder) has to die (/ death) as a consequence			
		to kill (/a murder) leads (someone) to die (/ death)			
	b	*to die / (a) death causes to kill			
		*to die / (a) death has to kill as a consequence			
		*to die / (a) death leads (someone) to kill			
Effect:		{to kill} (X)	CAUSES	{to die} (Y)	factive
		{to die} (Y)	IS_CAUSED_BY	{to kill} (X)	reversed
		{to kill}	CAUSES	{death}	factive
		{death}	IS_CAUSED_BY	{to kill}	reversed
		{murder}	CAUSES	{to die}	factive
		{to die}	IS_CAUSED_BY	{murder}	reversed
		{murder}	CAUSES	{death}	factive
		{death}	IS_CAUSED_BY	{murder}	reversed

Obviously, the event of 'dying' is not necessarily caused by 'killing'. This may either follow from the fact that the verb *kill* is only one out of the possible disjunct causes for *die*, or it may be expressed by explicitly labeling "dying IS_CAUSED_BY killing" as reversed (as is done here).

The following test is for detecting factive causation relation between dynamic verbs/nouns and static adjectives/adverbs:

Test 37		**Factive causation relation between verbs and adjectives (or adverbs)**			
yes	a	X causes to be Y			
		X has being Y as a consequence			
		X leads to be(ing) Y			
no	b	the converse of (a)			
Conditions:		- X is a verb in the infinitive form			
		- Y is and adjective			
Example:	a	to kill causes to be dead			
		to kill has being dead as a consequence			
		to kill leads someone to be dead			
	b	*to be dead causes to kill			
		*to be dead has to kill as a consequence			
		*to be dead leads (someone) to kill			
Effect:		{to kill} (X)	CAUSES	{dead} (Y)	factive
		{dead} (Y)	IS_CAUSED_BY	{to kill} (X)	reversed

Non-factivity is elicited with modal auxiliaries:

Test 38		**Non-factive causation relation between verbs/nouns using a modal auxiliary**		
yes	a	(A/an) X may cause (a/an) Y (A/an) X may have (a/an) Y as a consequence (A/an) X may lead to (a/an) Y		
no	b	the converse of (a)		
Conditions:		- X is a verb in the infinitive form or X is a noun in the singular - Y is a verb in the infinitive form or Y is a noun in the singular		
Example:	a	to search may cause to find to search may have to find as a consequence to search may lead (someone) to find		
	b	?to find may cause to search ?to find may have to search as a consequence ?to find may lead (someone) to search		
Effect:		{to search} (X)	CAUSES	{to find} (Y)(non-factive)
		{to find} (X	IS_CAUSED_BY	{to search} (Y)(non-factive)

The above tests are general tests to identify causal relation. More specific tests to elicit the different temporal relations of the situations. The following test elicits a ‘genuine’ cause relation between disjoint situations:

Test 39		**Causation relation between verbs/nouns referring to temporally disjoint situations**			
yes	a	If (a/an) X takes place it causes/may cause (a/an) Y to take place afterwards/later on			
no	b	the converse of (a)			
Conditions:		- X is a verb in the gerundive form or X is a noun in the singular - Y is a verb in the gerundive form or Y is a noun in the singular			
Example:	a	If sending takes place it causes receiving to take place later on			
	b	* If receiving takes place it causes sending to take place later on			
Effect:		{to send} (X)	CAUSES	{to receive} (Y)	factive
		{to receive} (Y)	IS_CAUSED_BY	{to send} (X)	reversed

The next test elicits a causal relation between temporally overlapping situations:

Test 40		**Causation relation between verbs/nouns referring to temporally non-disjoint situations**			
yes	a	If (a/an) X takes place it causes/may cause (a/an) Y to take place at the same time			
no	b	the converse of (a)			
Conditions:		- X is a verb in the gerundive form or X is a noun in the singular - Y is a verb in the gerundive form or Y is a noun in the singular - X and Y are not connected by means of the hyponymy relation			
Example:	a	If pulling takes place it may cause opening to take place at the same time			
	b	? If opening takes place it may cause pulling to take place at the same time			
Effect:		{to pull} (X)	CAUSES	{to open} (Y)	(non-factive)
		{to open} (Y)	IS_CAUSED_BY	{to pull} (X)	(non-factive)

As explained above, if two words only pass the above test, they should also be tested for a hyponymy relation.

Finally, we have stated that dynamic situations may cause other dynamic or non-dynamic situations. Dynamicity of the result can be inferred from the relation with a dynamic/non-dynamic hyperonyms (e.g. *state* or *change*). For example:

i)	fall asleep V	CAUSES	sleep V, sleep N, asleep A
	fall asleep V	HAS_HYPERONYM	change V
	sleep V	HAS_HYPERONYM	be V
	sleep N, asleep A	XPOS_NEAR_SYNONYM	sleep V

ii)	addormentare (make sleep) V	CAUSES	addormentarsi (fall asleep) V
	addormentarsi (fall asleep) V	HAS_HYPERONYM	cambiare (change) V
	addormentare (make sleep) V	HAS_HYPERONYM	fare (make, cause) V

In i) we see that the CAUSED verb *to sleep* is non-dynamic, as is expressed by its hyponymy relation with the verb *to be*. We also see that the noun *sleep* and the adjective *asleep* have near-synonymy relations with it and must, therefore, also be non-dynamic. In ii) we see an example in which the Italian verb *addormentarsi* (to fall asleep) is caused by *addormentare* (to make sleep). The fact that we are dealing with two dynamic situations is again expressed by the hyponymy relation: *addormentarsi* is a 'non-controlled' process and *addormentare* is a 'controlled' action.

2.2.3.8. HAS_SUBEVENT and IS_SUBEVENT_OF

According to Fellbaum (Miller et al, 1990: 45) the entailment relation underlies all verbal relations: "the different relations that organize the verbs can be cast in terms of one overarching principle, lexical entailment". Next, lexical entailment is differentiated on the basis of the temporal relation between events and the direction of the implication or entailment:

a. + Temporal Inclusion (the two situations partially or totally overlap)
- a.1 co-extensiveness (e. g., *to limp/to walk*) hyponymy/troponymy
- a.2 proper inclusion (e.g., *to snore/to sleep*) entailment

b. - Temporal Exclusion (the two situations are temporally disjoint)
- b.1 backward presupposition (e.g., *to succeed/to try*) entailment
- b.2 cause (e.g., *to give/to have*)

In the actual database the relation Entailment is applied to those cases that cannot be expressed by the more specific hyponymy and cause relations. In that case at least the direction of the implication or entailment is indicated. In the case of *snore/sleep* the direction is from *snore* to *sleep*: i.e. *snore* implies *sleep* but not the other way around. In the case of *buy/pay* on the other hand *buy* implies *pay* but not the other way around.

In EuroWordNet, the differences in the direction of the entailment can however be expressed by the labels factive and reversed. For example, 'backward presupposition' can be expressed by using the causal relation in conjunction with the factivity label:

{to succeed}	IS_CAUSED_BY	{to try}	*factive*
{to try}	CAUSES	{to succeed}	*non-factive*

Fellbaum (1998) already suggests that the 'proper inclusion' is more intuitively described by a verb meronymy relation. She then abandons this solution because the entailment from "snore" to "sleep" is reversed compared to "buy" and "pay". However, such implicational differences can also occur for noun-meronyms: e.g. "car" implies "door" but "door" is not necessarily part of a "car", "propeller" is part of an "aircraft", but an "aircraft" does not necessarily have a "propeller". We have seen that this implicational difference is encoded by the label reversed. The same can be done for the above verbs in combination with a HAS_SUBEVENT/ IS_SUBEVENT_OF relation:

{to snore}	IS_SUBEVENT_OF	{to sleep}	
{to sleep}	HAS_SUBEVENT	{to snore}	*reversed*
{to buy}	HAS_SUBEVENT	{to pay}	
{to pay}	IS_SUBEVENT_OF	{to buy}	*reversed*

The SUBEVENT relation is very useful for many closely related verbs and appeals more directly to human-intuitions (parallel to part-whole relation of concrete entities).

In the following tests, general criteria for the definition of the HAS_SUBEVENT relation between verbs (/nouns referring to events or processes) are given:

Test 41		**Has_Subevent/Is_Subevent_of relation between verbs/nouns (a)**			
yes	a	Y takes place during or as a part of X, and whenever Y takes place, X takes place			
no	b	the converse of a)			
Conditions:		- X is a verb in the gerundive form - Y is a verb in the gerundive form			
Example:	a	Snoring takes place during or as part of sleeping, and whenever snoring takes place, sleeping takes place			
	b	*Sleeping takes place during or as part of snoring *Whenever sleeping takes place, snoring takes place			
Effect:		{to snore} (X)	IS_SUBEVENT_OF	{to sleep} (Y)	
		{to sleep} (Y)	HAS_SUBEVENT	{to snore} (X)	reversed

Test 42		**Has_Subevent/Is_Subevent_Of relation between verbs/nouns (b)**		
yes	a	X consists of Y and other events or processes		
no	b	the converse of a)		
Conditions:		- Y is a verb in the gerundive form - X is a verb in the gerundive form		
Example:	a	buying consists of paying and other events or processes		
	b	*paying consists of buying and other processes		
Effect:		{to buy} (Y)	HAS_SUBEVENT	{to pay} (X)
		{to pay} (X)	IS_SUBEVENT_OF	{to buy} (Y) reversed

2.2.3.9. IN_MANNER and MANNER_OF

The notion of troponymy in WordNet1.5 is motivated by manner-verbs (e.g. manners of movement) and their more general superordinate, e.g. “slurp” can paraphrased as “to eat noisely” and is encoded as a troponym of “eat”. Troponymy can be seen as a subtype of hyponymy: i.e. it implies hyponymy and a manner feature. Still, the trponnymy relation has been used to encode all hyponymy relation in the database, even in cases where the manner is not implied. In EuroWordNet, we decided not to differentiate between troponymy and hyponymy but to use the IN_MANNER and MANNER_OF relation in addition to normal hyponymy to make the manner component explicit (if it is significant in the meaning of the verb):

Test 43		**to take place in certain manner**		
yes	a	to X is to Y in a Z manner/way.		
Conditions:		X and Y are verbs Y is the hyperonym of X Z is an adjective/adverb		
Example:	a	to slurp is to eat in a noisely manner X = slurp, Y = eat Z = noisely		
Effect:		slurp V	HAS_HYPERONYM	eat V
		slurp V	IN_MANNER	noisely Adverb
		noisely Adverb	MANNER_OF	slurp V reversed

2.2.3.10. BE_IN_STATE and STATE_OF

This relation is needed to encode links between nouns that refer to anything in a particular state expressed by an adjective. These nouns often have an open denotation: i.e. they can refer to any entity to which the state applies, e.g. "the poor" refers to all entities which are in a "poor" state. Note that these nouns are not equivalent to the states: the entities that have the property "poor" are not states but normal 1stOrderEntities. This relation is therefore across different semantic types. The general test is:

Test 44		**being in a particular state**		
yes	a	a/an/the X is the one/that to whom/which the state Y applies		
Conditions:		X is a noun		
		Y is an adjective/adverb		
Example:	a	the poor are the ones to whom the state poor applies		
		X = poor N (a poor person)		
		Y = poor A		
Effect:		poor N	BE_IN_STATE	poor A
		poor A	STATE_OF	poor N reversed

2.2.3.11. Derivational relations

Two derivational relations have been taken over from WordNet1.5:

- DERIVED/ DERIVED_FROM/HAS_DERIVED
- PERTAINS_TO and IS_PERTAINED_TO

The DERIVED relation is a purely morphological relation. In addition to DERIVED there must also be some other semantic relation e.g. synonymy, antonymy, role, cause. The general relation DERIVED is used if it is not clear what is the base form and what form is derived.

The PERTAIN relation is a more specific morphological relation with an unclear semantic effect. It is used for many adjectives that can only be related to nouns as a kind of topic marker: atomic/atom, chemical/chemistry, Greek/Greece. The relation to the corresponding nouns can only be paraphrased as: concerning, related to. This relation is more vague than the previous relation because the adjective itself is meaningless. There is no positive test for this relation (except for related to) but it can be inferred from the fact that none of the other relations hold (causal, in_state) and the adjective itself is void. Obviously, the relation holds between variants only.

2.2.3.12. Instance and Class

Hyponymy is a relation between classes of entities. Individual entities can also be said to belong to some class. Although we do not find many instances in a lexical database, the relation is useful for users that want to add particular instances and do not want to consult a separate database. To distinguish it from hyponymy the relation is dubbed has_instance and its inverse belongs_to_class:

Test 45		**Individuals belonging to a class**
yes	a	X is one of the Ys
no	b	Y is one of the Xs
Conditions:		X is a proper noun
		Y is a noun
Example:	a	Manchester is one of the cities
Effect:		Manchester BELONGS_TO_CLASS city
		city HAS_INSTANCE Manchester

2.2.3.13. Undefined Relations: "fuzzynyms"

Finally, there is a relation to cover all the cases in which a word is strongly associated with another word but no proper relation has been defined. Fuzzynymy holds when all the above tests fail but the test *X has some strong relation to Y* still works. A FUZZYNYM relation holds between words with the same part-of-speech, XPOS_FUZZYNYM holds across part-of-speech.

2.3. Multilinguality

2.3.1 Equivalence relations

The Equivalence Relations between synsets in each language and the Inter-Lingual-Index are to a large extent parallel to the Language Internal Relations.

Table 4: The Equivalence Relations in EuroWordNet

EQ_RELATION	*Source Synsets*	*Target ILIs*
EQ_SYNONYM	diventare IT	to become
EQ_NEAR_SYNONYM	schoonmaken NL	to clean in X senses
EQ_HAS_HYPERONYM	kunstproduct NL (artifact substance)	artifact; product
EQ_HAS_HYPONYM	dedo ES (a finger or toe)	toe; finger
OTHER RELATIONS		
EQ_HAS_HOLONYM	EQ_IN_MANNER	EQ_BE_IN_STATE
EQ_HAS_MERONYM	EQ_CAUSES	EQ_IS_STATE_OF
EQ_INVOLVED	EQ_IS_CAUSED_BY	EQ_GENERALIZATION
EQ_ROLE	EQ_HAS_SUBEVENT	EQ_METONYM
EQ_CO_ROLE	EQ_IS_SUBEVENT_OF	EQ_DIATHESIS

The most important relation is EQ_SYNONYM, which only holds if there is a 1-to-1 mapping between synsets. In addition there are relations for complex-equivalence relations, among which the most important are:

- EQ_NEAR_SYNONYM when a meaning matches multiple ILI-records simultaneously, when multiple synsets match with the same ILI-record, or when there is some doubt about the precise mapping.
- EQ_HAS_HYPERONYM when a meaning is more specific than any available ILI-record.
- EQ_HAS_HYPONYM when a meaning can only be linked to more specific ILI-records.

The complex-equivalence relations are comparable to the kinds of mismatches across word meanings described in the Acquilex project in the form of complex *TLINKS* (Ageno et al 1993, Copestake et al. 1995, and Copestake and Sanfilippo 1993). It is possible to manually encode these relations directly in the database, but they can also be extracted semi-automatically using the technology developed in Acquilex. The difference between Acquilex and EuroWordNet is that the *TLINKS* in Acquilex are lexical transfer links between language-pairs at a sense-level, whereas the equivalence relations in EuroWordNet are established at the synset level from each language to a single interlingua (the ILI). Language-to-language mappings can only indirectly be inferred via the ILI.

In EuroWordNet, the complex relations are needed to help the relation assignment during the development process when there is a lexical gap in one language or when meanings do not exactly fit. The first situation, in which a single synset matches several ILI-records simultaneously, occurs quite often. The main reason for this is that the sense-differentiation in WordNet1.5 is more fine-grained than in the traditional resources from which the other wordnets are built. For example, in the Dutch resource there is only one sense for *schoonmaken* (to clean) which simultaneously matches with at least 4 senses of *clean* in WordNet1.5:

- {make clean by removing dirt, filth, or unwanted substances from}
- {remove unwanted substances from, such as feathers or pits, as of chickens or fruit}
- {remove in making clean; "Clean the spots off the rug"}
- {remove unwanted substances from - (as in chemistry)}

The Dutch synset *schoonmaken* will thus be linked with an EQ_NEAR_SYNONYM relation to all these senses of *clean*. A similar situation may arise when there is under-differentiation in the Dutch wordnet. For example, *keuze* in the Dutch resource is defined as the *act* or *result* of choosing, likewise it can be linked with EQ_NEAR_SYNONYM relations to both *choice*#1 (the act of choosing) and *choice*#2 (what is chosen) in WordNet 1.5.

Despite the sense-differentiation in WordNet1.5, the reverse situation also occurs. For example, *versiersel* and *versiering* are not coded as synonyms in the Dutch resource but they can still both be linked to the same WN1.5 synset *decoration*. It may be the case that the Dutch words should be merged into a single synset, but, they can also be related by a weaker NEAR_SYNONYM relation. In the latter case, they can share the same ILI-record but the equivalence relation should be EQ_NEAR_SYNONYM and not EQ_SYNONYM.

The EQ_HAS_HYPERONYM is typically used for gaps in WordNet1.5 or in English. Such gaps can be cultural or pragmatic. A cultural gap is a concept not **known** in the English/American culture, e.g. the Dutch noun *citroenjenever*, which is a kind of gin made out of lemon skin, or the Dutch verb: *klunen* (to walk on skates over land from one frozen water to another). Pragmatic gaps are caused by lexicalization differences between languages, in the sense that in this case the concept is known but not expressed by a single lexicalized form in English., e.g.:

Dutch: *doodschoppen* (to kick to death),
Spanish: *alevín* (young fish),
Italian: *rincasare* (to go back home).

In these cases the lexicalization patterns in the languages are different from English but the concepts are familiar to all cultures. Typically, a concept like "doodschoppen" (kick to death) in Dutch will get two eq_hyperonym relations, one to "to kill" and one to "to kick". This is parallel to the multiple hyperonyms the word will receive in Dutch. Similarly, Spanish "alevín" (young fish) can both be linked with an eq_hyperonym to "fish" and eq_be_in_state to "young". Using multiple equivalence relations the meanings of some synsets can be exhaustively linked to the ILI.

In all the above cases, the non-English word is more specific and thus can be related to a more general English ILI-concept using an EQ_HAS_HYPERONYM relation. The EQ_HAS_HYPONYM is then used for the reversed situation, when WordNet1.5 only provides more narrow terms. An example is Spanish *dedo* which can be used to refer to both *finger* and *toe*. In this case there can only be a pragmatic difference, not a genuine cultural gap.

A special case of gaps are mismatches in Part of Speech across languages, e.g. in Dutch the adjective *aardig* is equivalent to the verb *to like* in English but there is no verb with that meaning in Dutch. The equivalence relations to the ILI are however not sensitive to the Part-of-Speech. It is thus possible to directly express an EQ_NEAR_SYNONYM relation between *aardig* Adjective and *like* Verb.

The complex equivalence relations are expressed separately from each language to the index. Decisions on the matching are taken by each site separately for their language, towards the English ILI. In addition, there is also an effort to smoothen the matching across the wordnets by adapting the index. This will be discussed in the next subsection.

2.3.2. Inter-Lingual-Index

As explained in the introduction, the Inter-Lingual-Index (ILI) is an unstructured fund of concepts, with the only purpose to provide an efficient mapping across languages. Each concept is represented as a ILI-record that in principle consists of a synset, a part-of-peech label, a gloss and a reference to its source. The ILI started off as a plain list of WordNet1.5 synsets, but it has been adapted to provide a better matching across the wordnets. There are several changes to the WordNet1.5 list of concepts:

- adding missing concepts occurring in other wordnets
- creating more global sense clusterings
- to add domain terminology for computing terms
- improve the glosses

In 425 cases, a missing gloss was manually added to an ILI-record derived from WordNet1.5. Glosses are often crucial for determining proper equivalence relations. The other changes are discussed in the next subsections.

2.3.2.1. Extending the ILI with new concepts

First of all, there are concepts in the local wordnets which are not present in WordNet1.5, e.g. *a female cashier*. To be able to still express equivalence relations between such a concept in other wordnets (*cajera* in Spanish, *cassière* in Dutch), the ILI has to be extended. The ultimate ILI will thus become the superset of all concepts occurring in 2 or more wordnets. The procedure for extending the ILI is as follows. All sites send descriptions of the gaps in the form of potential new ILI-records to one site. The ILI-records are described using a formalized semantic specification so that the candidates can be compared. If there is sufficient overlap between at least two descriptions, a new ILI-record is added and the local synsets referring to this new ILI-record will get an additional EQ_SYNONYM relation to this record. These synsets will thus have at least two different equivalence relations, a complex equivalence relation to the closest WordNet1.5 synset and a simple equivalence relation to the new ILI-record, e.g.:

Spanish Wordnet		**ILI**		**Dutch Wordnet**
cajera	eq_hyperonym	{*cashier}*	eq_hyperonym	*cassière*
	eq_synonym	{ *female cashier*}	eq_synonym	

This example shows that it is possible to extract direct equivalences in Dutch and Spanish, but also to find the closest matches with English (albeit a more specific concept). Due to lack of time and resources in the project, we have not been able to actually extend the ILI with new concepts, based on evidence from other wordnets. Furthermore, the discussion about the different status of mismatches to the ILI is still ongoing (see Vossen, Peters, Gonzalo 1999, for a further discussion).

Nevertheless, the ILI has been extended with computer terminology to illustrate the possibility of incorportating domain terminology in the generic wordnets. In total, 444 ILI-records have been labelled as Computer Terminology. The selection has been based on a number of electronic resources:

- FOLDOC Free On-line Dictionary of Computing: http://wombat.doc.ic.ac.uk/foldoc/index.html. Around 6000 entries with definitions and subdomain information.
- DATA Direct glossary: http://data-direct.com/glossary.htm, around 650 entries with definitions
- Dartek glossary: http://www.dartek.com/glossary/glossary.cfm, around 1000 entries with definitions
- Netglos glossary: http://wwli/com/translation/netglos/netglos.html, around 110 entries with definitions

These terms have been verified using a word frequency list taken from:

- Ami-Pro manual from Lotus (Donker, Serail and Vossen 1994)
- the Brittish National Corpus
- Unix manuals

The selected terms have been matched against the WordNet1.5 vocabulary. If the concepts where present in WordNet1.5 in the correct sense, the corresponding ILI-record has been labelled as computer term by adding a domain label "**COMPUTER_TERMINOLOGY**" to the gloss. This happened for 107 concepts. In the other cases, 337 concepts, we added new ILI-records to the ILI (with the appropriate synset, gloss, and part-of-speech). In total, 397 nouns, 32 verbs and 15 adjectives have been added.

2.3.2.2. Creating a coarser level of differentiation in the ILI

Even though the ILI should ideally be the superset of concepts occurring in the different wordnets, it should, on the other hand, not be too fine-grained either. If many subtle senses are distinguished, it is more complicated to establish equivalences across the wordnets. In the case of "clean", for example, it may be that different sites link equivalent synsets to different meanings, resulting in a mismatch across the languages. A similar mismatch may be caused by inconsistent enumeration of regular polysemy across resources. In the ILI, there are different synsets for *university* as a building and *university* as the organization, and in fact many institute/building pairs are present. However, in other wordnets we may find situations where only one of the senses is given. If a different choice is made for the *building* or the *institute*, synsets cannot be matched across wordnets. The second adaptation to the ILI therefore aims at grouping senses that can be related by 'regular polysemy' (Apresjan 1973; Copestake and Briscoe 1991; Nunberg and Zaenen 1992). This is achieved by adding so-called Composite ILI-records, which can be compared with Complex Types as defined by Pustejovsky (1995).

For example, the synsets in Dutch, Spanish and Italian in the next table are related via EQ_SYNONYM or EQ_NEAR_SYNONYM relations to ILI-records that represent 5 different senses of "office": place; actions carried out; job; organization and the group of people. The synsets are separated by curled brackets. In some cases multiple synsets are linked to the same ILI-record.

Table 5: Dutch, Spanish and Italian Synsets linked to senses of "office" in the ILI.

ILI record	Dutch Synsets	Spanish Synsets	Italian Synsets
{**office**}-1960921 where professional or clerical duties are performed; "he rented an office in the new building"	{kantoor; werkkamer; werkruimte}	{oficina}	{ufficio; studio}
{role; part; **office**; function}-399406 the actions and activities assigned to or required or expected of a person or group: "the function of a teacher";"the government must do its part" or "play its role" or "do its duty"	{functie; rol} {emplooi}	{función; papel; officio}	{ufficio; mansione; carica}
{situation; place; spot; **office**; slot; berth; post; position}-344376 a job in an organization or hierarchy; "he ocupied a post in the treasury"	{ambt; ambtsbediening; bediening; officie; officium} {betrekking; baan; dienstbetrekking; dienstverband; functie; job; positie; werk; werkkring} {arbeidsplaats; plaats}	{caro; puesto}	{lavoro; impiego; occupazione}
{authority; **office**; bureau; agency}-5301461 an administrative unit of government; "the Central Intelligence Agency"; "the Census Bureau"; "Office of Management and Budget"; "Tennessee Valley Authority"	{dienst} {kantoor; bureau; bureel; burelen} {bureau} {agentuur}	{agencia; oficina}	{ispettorato}
{office staff; **office**}-5303509 professional or clerical workers in an office; "the whole office was late the morning of the blizzard"	{kantoorpersoneel}		

Several things can be observed here. First of all, we see that the polysemy is not parallel across the languages. In the Spanish wordnet, only "oficina" is polysemous relative to "office" and in the Italian and Dutch wordnet only "ufficio" and "kantoor" are, respectively. Furthermore, each of these is polymous over different senses of "office" and only maps to 2 out of the 5 senses (obviously, many of these words may be polysemous in other senses not related to "office" in English). In most cases, the concepts are lexicalized by different forms, derivations or compounds. Finally, we see that {office staff; **office**}-5303509 is only represented in Dutch.

A native speaker of Spanish and Italian has to confirm whether variants in the synsets in Spanish and Italian related to "office" can take the meaning of "the group of people working in an office". This is definitely the case for some of the Dutch variants: "dienst", "kantoor", bureau", "werk". Apparently, the polysemy in the wordnets is more parallel then the direct linking suggests. The resources used to build the wordnets have not been consistent in explicating all the different senses.

By creating a grouping for all these senses of "office" in EuroWordNet, we can still establish this potential relation. Such a grouping is made by the next example of a Composite ILI-record for "office" that relates the 5 senses by a metonymy relation. The example is in the ILI-import format that will be explained later in section 2.4. This ILI-record establishes a grouping of the senses listed as variants via the EQ_RELATION to the target concepts. The target concepts are represented by the `WORDNET_OFFSET` numbers:

```
0 ILI_RECORD
      1 PART_OF_SPEECH "n"
      1 ADD_ON_ID 20
               1 GLOSS "a job in an organization or hierarchy; "he ocupied a post
               in the treasury""\building where professionals work or the
               institution represented by these professionals"\professional or
               clerical workers in an office; "the whole office was late the
               morning of the blizzard""\the actions and activities assigned to or
               required or expected of a person or group: "the function of a
               teacher"; "the government must do its part" or "play its role" or
               "do its duty""
      1 VARIANTS
            2 LITERAL "office"
                  3 SENSE 1
            2 LITERAL "office"
                  3 SENSE 2
            2 LITERAL "office"
                  3 SENSE 4
            2 LITERAL "office"
                  3 SENSE 5
            2 LITERAL "office"
                  3 SENSE 6
      1 EQ_RELATION "eq_metonym"
            2 TARGET_ILI
                  3 WORDNET_OFFSET 1960921
                  3 WORDNET_OFFSET 344376
                  3 WORDNET_OFFSET 399406
                  3 WORDNET_OFFSET 5301461
                  3 WORDNET_OFFSET 5303509
```

Whenever such a Composite ILI-record is added to the ILI, the EuroWordNet database will automatically generate additional equivalence relations for all synsets in the wordnets related with an EQ_SYNONYM or EQ_NEAR_SYNONYM relation to any of the specific meanings that are grouped by this ILI-record. All the synsets in the above table will thus receive an additional eq_metonym link to the Composite ILI-record, as is shown in the next figure for "oficina" in Spanish:

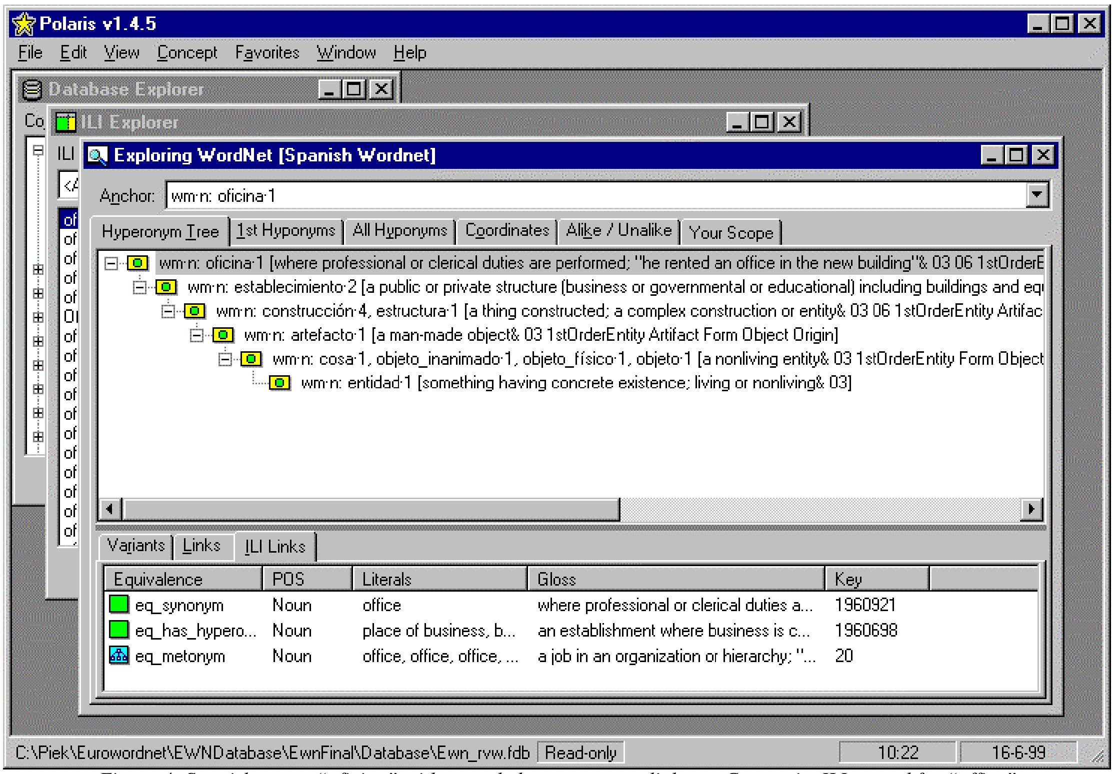


Figure 4: Spanish synset "oficina" with extended EQ_METONYM link to a Composite ILI-record for "office"

Even though, none of the local wordnets has the same differentiation, all synsets now share the metonymy link and, likewise, can be retrieved in a global way when we look for synsets to the same ILI-record with EQ_METONYM. This can either be used to extend the wordnets with new senses for the words in these synsets or link the synset to new words. Alternatively, the database can be used in a more global way to expand synsets across languages via EQ_METONYM relations, even thouh this might overgenerate.

Similar Composite ILI-records are added for generalizations that group over-differentiation as we have seen for "clean" (related by EQ_GENERALIZATION) and for enumerated senses that reflect diathesis alternations for verbs (related by EQ_DIATHESIS), such as between causative and inchoative pairs, e.g.:

hit 1: hit a ball (synonym: cause to move by striking)
hit 2: come into sudden contact with: "The arrow hit the target"
hit 3: deal a blow to; "He hit her hard in the face"

Differences in arity and the semantic characterization of subcategorized arguments highlight different perspectives on the situation described by the predications, or express semantic notions such as 'causation' and 'result of causation' (Levin 1993). By relating these diathesis alternation patterns to more Composite ILI-records we will thus be able to link local synsets regardless of whether the verbs in question display dissimilar alternation patterns in different senses, have a number of alternations collapsed in a single sense, or are monosemous.

Buitelaar (1998), Peters et al. (1998) describe how these sense-groups can be extracted from a resource such as WordNet1.5. Peters (1999) gives a complete description of the extracted Composite ILI clusters in EuroWordNet. Here we give just an overview:

Table 6: Composite ILI-records

	Metonymy			*Generalization*		
	Clusters	Words	Senses	Clusters	Words	Senses
nouns	30	24	67	1703	1398	3205
verbs	0	0	0	2905	1799	5134

The clusters have been derived according to the following methodologies:

- manual clustering (generalization and metonymy)
- automatically derived clusters (generalization)
    - based on the internal structure of Wn1.5 (sisters, autohyponymy)
    - based on matching WN15 with other resources
        - Levin's semantic classes underlying diathesis alternations (Levin 1993)
        - WN1.6
        - around 66 clusters based on one to many links between Dutch and Italian wordnets to the ILI
- 10 regular polysemy patterns derived from sense distribution in WN15 (e.g. 'music - dance', 'container - collection')

The sense-groupings lead to a more coarse differentiation of senses which will make the ILI more effective for mapping senses across languages. Inconsistency of sense-differentiation, such as for synsets related to *office*, will be captured by metonymy classes.

2.3.3. Accessing complex equivalence mappings

From what has been said so far it follows that there can be many-to-many mappings from local synsets to ILI-records. This may either be an EQ_NEAR_SYNONYM relation from and/or to multiple synsets (possibly with different part-of-speech), or an EQ_HAS_HYPONYM/ EQ_HAS_HYPERONYM and an EQ_SYNONYM to a new ILI-record, or various combinations of these (or other types of equivalence relations). Finally, it is possible that a single synset in a wordnet is linked to both a Composite ILI-record with an EQ_METONYM, EQ_DIATHESIS or EQ_GENERALIZATION and to one of the more specific senses grouped by the Composite ILI.

Table 7: Overview of mapping relations to the ILI

Relation	POS	Source Synsets : Target ILIs	Example
eq_synonym	same	1:1	auto : car
eq_near_synonym	any	many : many	apparaat, machine, toestel : apparatus, machine, device
eq_hyperonym	same	many : 1 (usually)	citroenjenever: gin
eq_hyponym	same	(usually) 1 : many	dedo : toe, finger
eq_metonymy	same	many/1 : 1	universiteit, universiteitsgebouw: university
eq_diathesis	same	many/1 : 1	raken (cause), raken: hit
eq_generalization	same	many/1 : 1	schoonmaken : clean

Note that a many-to-many mapping from a wordnet to the ILI, may also cause a further spreading when multiple ILI-records are next mapped to another wordnet. In the next screen-dump we see how such a fuzzy mapping results for *machine*, *apparatus*, *tool* in Dutch and Italian. In this example, 3 near synonyms in the Dutch wordnet are linked to multiple ILI-records, from-top-to-bottom: *device*, *apparatus*, *instrument*, *implement*, *tool*. The ILI-records are again represented by their glosses, where the synset of the highlighted ILI-record (device:1) is shown in the small box at the bottom-right corner. In the Italian wordnet we see that 4 of these ILI-records are given as EQ_NEAR_SYNONYMs of a single synset *utensile:1* but *device* is linked to *ferrovecchio:2* by an EQ_HAS_HYPERONYM relation (as indicated by the symbols).

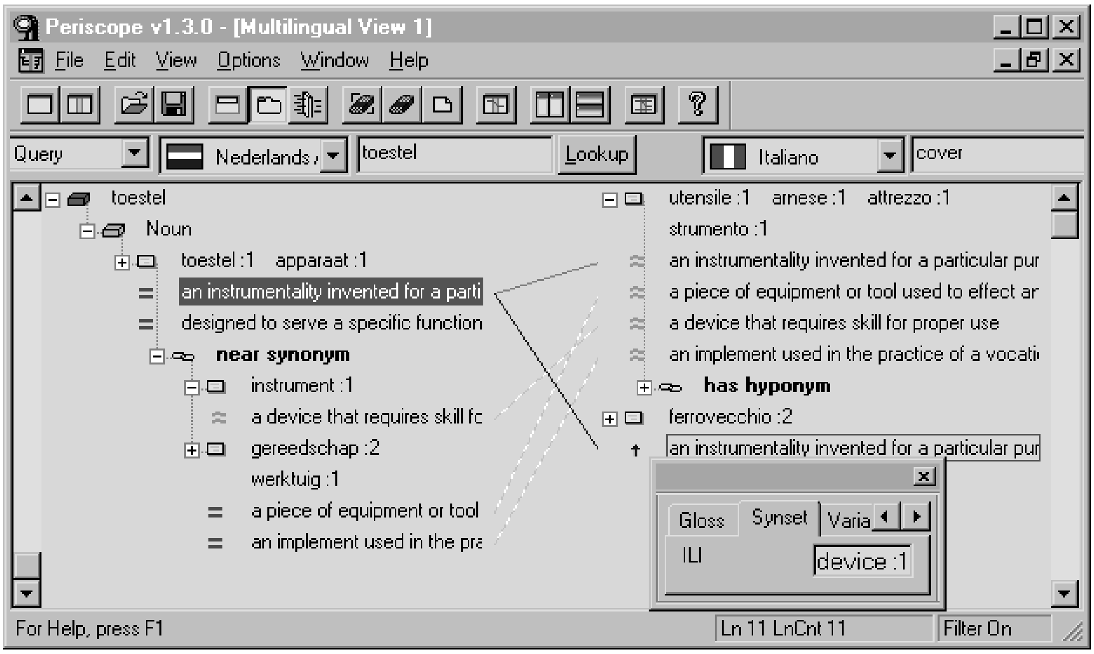


Figure 5: Many-to-many mappings of near synonyms of "apparatus" *synsets to ILI-records.*

Another important characteristic of the equivalence relations is the fact that they are established at the synset level. This is different from a traditional bilingual dictionary where specific relations are expressed between individual words or word-senses. For example, a pejorative term such as "idiot" is usually translated in a bilingual dictionary by a pejorative term in a the target language. In EuroWordNet, both the pejorative and the neutral term are members of the same synset and may have a single ILI-record as equivalent. Finally, the POS of an ILI-record is not relevant for creating equivalence links, e.g.: a nominal synset can have equivalence links to verbal and adjectival ILI-records, although the type of equivalence should be eq_near_synonym.

In general, we can thus say that the effect of the multilingual relations in EuroWordNet is that concepts are matched rather than words, that multiple concepts may share ILI-records (index-terms) or single concepts may yield multiple ILI-records. Furthermore, the ILI may be accessed very specifically by EQ_SYNONYM relations only, or by indicating any of the other complex equivalence mappings. The database thus provides the possibility to project a single concept or a cluster of concepts to another language, either specifically or in a more fuzzy way.

Once we have accessed a cluster of concepts in the target language, we can further use the language-internal relations to see the conceptual dependencies between these words (and possibly other words). This may point to solutions for gaps in the target language as is illustrated in Figure 6, where Dutch compound verbs for *ways of killing* are not lexicalized in English.

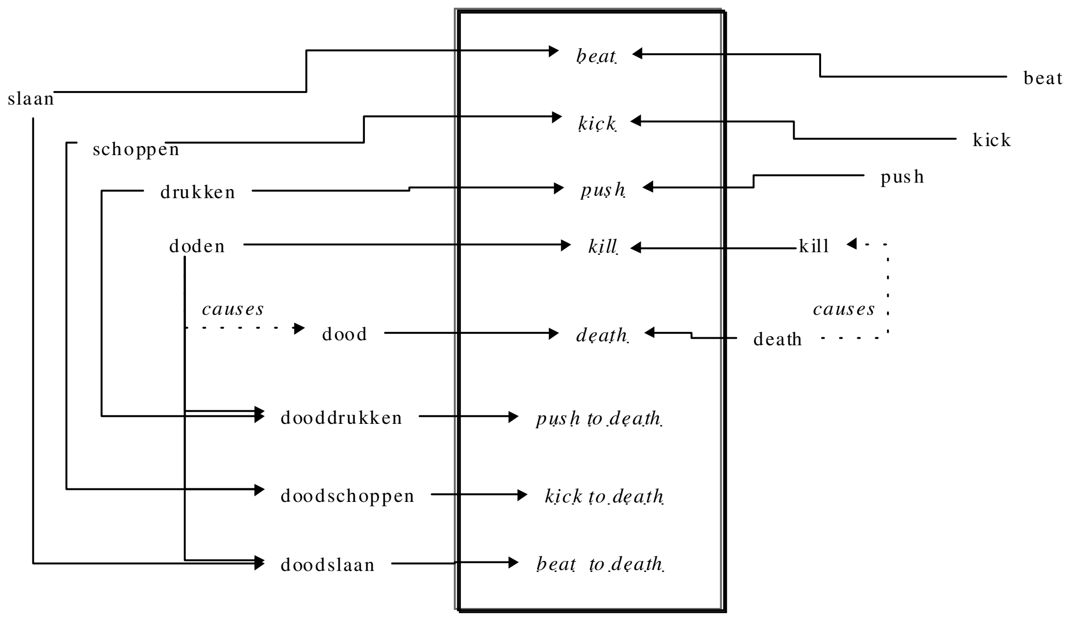


Figure 6: Ways of "killing" *lexicalized in Dutch and not in English.*

Here we see that the ILI is extended to represent concepts for the Dutch verbs, and there is no mapping to English verbs at the right side. The Dutch verbs have multiple hyperonyms to both the manner in which the event takes place (*beat*, *kick*, *push*) and the result (*kill*). Furthermore, *doden* and *kill*, which are equivalents, have a causal relation to the nouns *dood* and *death*, which are equivalent too. From this we may develop a strategy to generate expressions such as "kill by kicking" or "kick to death" as equivalents for the Dutch verb "doodschoppen".

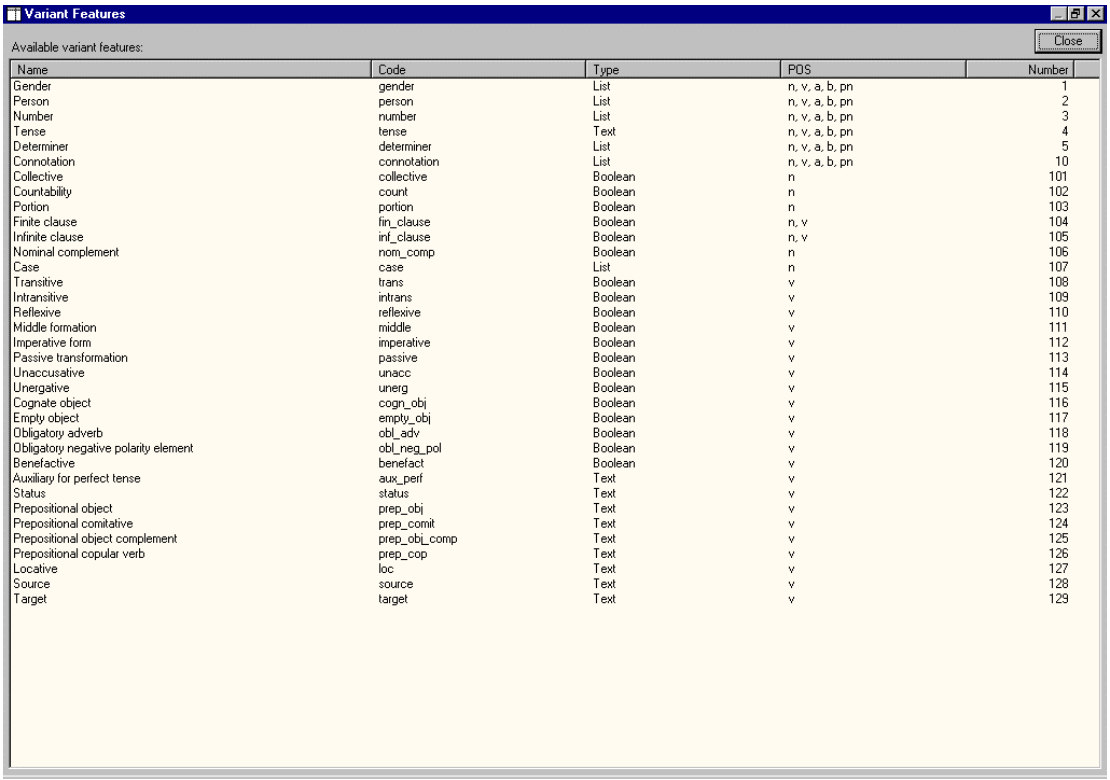
Concluding, we can say that instead of a single or a few specific alternatives in a bilingual dictionary, the EuroWordNet database gives a more comprehensive overview of concept-lexicalization in the target language, from which to choose the best candidate. In this sense, we can make a parallel with the 'Shake and Bake' methodology in Machine Translation (Whitelock 1992), where first an abstraction is made from the structural properties in the Source Language to a more neutral conceptual level (Shake), and next a (possibly different) new structure is generated in the target language (Bake). In the case of EuroWordNet, we are dealing with *lexical Shake*: abstract from the lexicalization that may be specific for a language (Vossen 1999). Bake is then possible by selecting the most appropriate candidate on the basis of co-occurrence restrictions in the target language, or the pragmatic and morpho-syntactic properties of the members in the synset. This kind of information can be extracted from Parole lexicons properly linked to the EuroWordNet database (see also Dorr et al. 1998).

2.4. Variant Information

For each variant in the synset specific information can be provided:

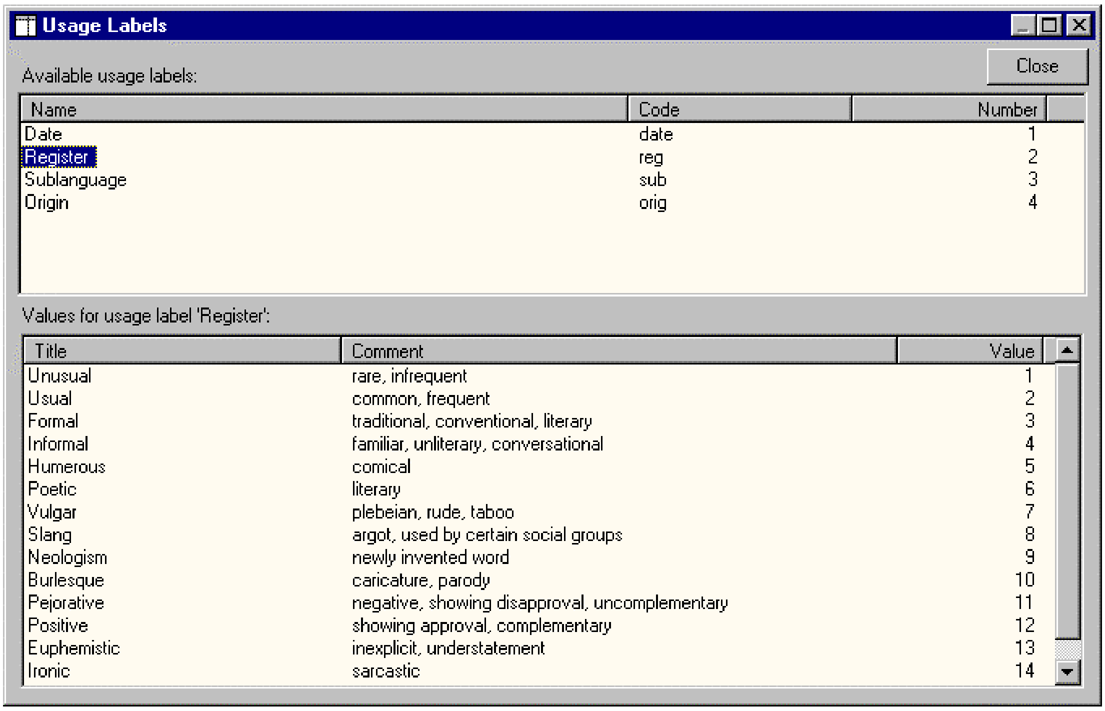
Usage Labels	Features on register, style, sub-domains.
Features	Morpho-syntactic properties for each part-of-speech.
Examples	Example sentences.
Translations	Whereas EuroWordNet provides equivalence links at the synset level, it is possible to specify here translations at the variant level.
Corpus Refs	Corpus reference information for the variant and corpus frequency
Data Source Refs	Data source reference information for the variant.
Definition	A single definition per variant
Status	Any label providing a status indication.
Parole ID	A reference to a specific Parole entry

Most of this information is optional. Builders of the wordnet are free to specify the examples, translations, corpus and data source references, the definition and the status. The Usage Labels and the Features have been defined more specifically, as is indicated in Figure 7 and 8.



Name	Code	Type	POS	Number
Gender	gender	List	n, v, a, b, pn	1
Person	person	List	n, v, a, b, pn	2
Number	number	List	n, v, a, b, pn	3
Tense	tense	Text	n, v, a, b, pn	4
Determiner	determiner	List	n, v, a, b, pn	5
Connotation	connotation	List	n, v, a, b, pn	10
Collective	collective	Boolean	n	101
Countability	count	Boolean	n	102
Portion	portion	Boolean	n	103
Finite clause	fin_clause	Boolean	n, v	104
Infinite clause	inf_clause	Boolean	n, v	105
Nominal complement	nom_comp	Boolean	n	106
Case	case	List	n	107
Transitive	trans	Boolean	v	108
Intransitive	intrans	Boolean	v	109
Reflexive	reflexive	Boolean	v	110
Middle formation	middle	Boolean	v	111
Imperative form	imperative	Boolean	v	112
Passive transformation	passive	Boolean	v	113
Unaccusative	unacc	Boolean	v	114
Unergative	unerg	Boolean	v	115
Cognate object	cogn_obj	Boolean	v	116
Empty object	empty_obj	Boolean	v	117
Obligatory adverb	obl_adv	Boolean	v	118
Obligatory negative polarity element	obl_neg_pol	Boolean	v	119
Benefactive	benefact	Boolean	v	120
Auxiliary for perfect tense	aux_perf	Text	v	121
Status	status	Text	v	122
Prepositional object	prep_obj	Text	v	123
Prepositional comitative	prep_comit	Text	v	124
Prepositional object complement	prep_obj_comp	Text	v	125
Prepositional copular verb	prep_cop	Text	v	126
Locative	loc	Text	v	127
Source	source	Text	v	128
Target	target	Text	v	129

Figure 7: Morpho-syntactic variant features allowed in EuroWordNet



Usage Labels

Close

Available usage labels:

Name	Code	Number
Date	date	1
Register	reg	2
Sublanguage	sub	3
Origin	orig	4

Values for usage label 'Register':

Title	Comment	Value
Unusual	rare, infrequent	1
Usual	common, frequent	2
Formal	traditional, conventional, literary	3
Informal	familiar, unliterary, conversational	4
Humerous	comical	5
Poetic	literary	6
Vulgar	plebeian, rude, taboo	7
Slang	argot, used by certain social groups	8
Neologism	newly invented word	9
Burlesque	caricature, parody	10
Pejorative	negative, showing disapproval, uncomplementary	11
Positive	showing approval, complementary	12
Euphemistic	inexplicit, understatement	13
Ironic	sarcastic	14

Figure 8: Usage labels for variants allowed in EuroWordNet

As said before, most of these features are optional. They have been used during the building process. In Appendix IV, we give the allowed variant features and their values. For further details on the labels and fields that can be stored in the database, we refer to the Polaris user manual (Louw 1998, D024) that can be downloaded form the EuroWordNet WEB site.

2.5. EuroWordNet Import/Export Format

The EuroWordNet data are distributed as a database and as plain text files. The text files are structured according to the EuroWordNet import/export format. This is the format that the Polaris database (see section 4 below) can read and will generate when concepts are exported. There are 3 different formats:

- Synsets
- ILI-records
- Top-Concepts and Domains

2.5.1. Import/Export format for synsets

The synset format is used for importing concepts for a language-specific wordnet. All the distributed wordnets are delivered in this format. Below is a (nonsensical) made-up example of a synset structure in the import format, illustrating many options:

```
0 @55718@ WORD_MEANING
  1 PART_OF_SPEECH "n"
  1 VARIANTS
    2 LITERAL "job"
      3 SENSE 2
      3 DEFINITION "what you should do for a living"
      3 EXTERNAL_INFO
        4 SOURCE_ID 1
          5 TEXT_KEY "08508615-n"
   2 LITERAL "work"
     3 SENSE 1
     3 STATUS "New"
     3 DEFINITION "what you do for a living"
```

```
    3 USAGE_LABELS
      4 USAGE_LABEL "sub"
        5 USAGE_LABEL_VALUE "Medicine"
      4 USAGE_LABEL "reg"
        5 USAGE_LABEL_VALUE "Informal"
      4 USAGE_LABEL "orig"
        5 USAGE_LABEL_VALUE "Latin"
    3 FEATURES
      4 FEATURE "connotation"
        5 FEATURE_VALUE "figurative"
      4 FEATURE "gender"
        5 FEATURE_VALUE "feminine"
      4 FEATURE "number"
        5 FEATURE_VALUE "singular"
    3 EXTERNAL_INFO
      4 CORPUS_ID 2
        5 FREQUENCY 920575
      4 SOURCE_ID 1
        5 TEXT_KEY "II.6.a"
      4 SOURCE_ID 3
        5 NUMBER_KEY 8008
      4 PAROLE_ID 36721
1 INTERNAL_LINKS
  2 RELATION "has_hyponym"
    3 TARGET_CONCEPT
      4 PART_OF_SPEECH "n"
      4 LITERAL "lexicography"
        5 SENSE 9
 2 RELATION "has_hyperonym"
    3 TARGET_CONCEPT
      4 PART_OF_SPEECH "n"
      4 LITERAL "activity"
        5 SENSE 3
1 EQ_LINKS
  2 EQ_RELATION "eq_has_hyperonym"
    3 TARGET_ILI
      4 PART_OF_SPEECH "n"
      4 WORDNET_OFFSET 8508615
  2 EQ_RELATION "eq_near_synonym"
    3 TARGET_ILI
      4 PART_OF_SPEECH "n"
      4 WORDNET_OFFSET 2861550
  2 EQ_RELATION "eq_generalization"
    3 TARGET_ILI
      4 PART_OF_SPEECH "n"
      4 ADD_ON_ID 8543
```

The first line, starting with level "0" identifies the synset (called WORD_MEANING). If the synset is exported from a database, then a synset ID follows (@55718@). At the next level (1) information is given on:

- the part of speech: noun, verb, adjective, adverb or proper noun.
- the variants, synset members or synonyms
- the language-internal relations
- the equivalence relations

For each variant, the literal and sense are obligatory. Optionally information for each variant is given at level (3). The latter information includes the status (anything can be specified here), the usage labels and their values, morpho-syntactic features (FEATURES), and references to corpora and corpus frequency, pointers to sources and possible reference to a PAROLE entry. A full list of the optional variant features is provided in Appendix IV. The example also illustrates the different types of values: free-text, values, or numbers.

The language internal relations (INTERNAL_LINKS) are specified one by one by indicating the type of relation and the target concept. The target concept is indicated by the part-of-speech the literal and sense number of one of its variants. The equivalence relations (EQ_LINKS) follow a similar syntax, but the target is now an ILI-record either identified by the file offset position that originates from the original WordNet1.5 data file or, if the ILI record is added in EuroWordNet, a so-called ADD_ON id-number.

2.5.2. Import/Export format for ILI-records

The import format for ILI records follows a similar pattern as for synsets. The first lines identifies the record, the next levels contain the data. There are 3 subtypes of ILI-records:

- Simple ILI-records that originate from WordNet1.5
- Simple ILI-records that do not originate from WordNet1.5
- Composite ILI-records that represent a grouping of other ILI-records

ILI-records that originate from WordNet1.5 consist of a specification of the part-of-speech, a reference to the file offset position in the original WordNet1.5 database, the gloss and a list of variants representing the synset.

```
0 @1@ ILI_RECORD
  1 PART_OF_SPEECH "n"
  1 WORDNET_OFFSET 2403
  1 GLOSS "something having concrete existence; living or nonliving& 03"
  1 VARIANTS
    2 LITERAL "entity"
      3 SENSE 1

0 @2@ ILI_RECORD
  1 PART_OF_SPEECH "n"
  1 WORDNET_OFFSET 2728
  1 GLOSS "any living entity& 03 1stOrderEntity Living Natural Origin"
  1 VARIANTS
    2 LITERAL "life form"
      3 SENSE 1
    2 LITERAL "organism"
      3 SENSE 1
    2 LITERAL "being"
      3 SENSE 1
    2 LITERAL "living thing"
      3 SENSE 1
```

In some cases, glosses have been edited or added. This information is imported via a special kind of update format:

```
0 ILI_RECORD
        1 UPDATE
        1 PART_OF_SPEECH "n"
        1 FILE_OFFSET 8340478
        1 GLOSS "COMPUTER_TERMINOLOGY a unit of information (from Binary+digIT); the
amount of information in a system having two equiprobable states; "there are 8 bits in
a byte""
```

The second line here indicates the UPDATE function and the gloss will overwrite the gloss that is already in the database.

The second type of ILI format are the Composite ILI-records. The import records have a so-called ADD_ON_ID instead of an FILE_OFFSET number to identify the record. Furthermore, they have equivalence relations to the ILI-records that are grouped by it. For the rest the structure is the same:

```
0 ILI_RECORD
        1 PART_OF_SPEECH "v"
        1 ADD_ON_ID 3029
        1 GLOSS "give certain properties to something; "get someone mad"; "She made us
look silly"; "He made of fool of homself at the meeting"; "Don''t make this into a big
deal"; "This invention will make you a famous physicist""
        1 VARIANTS
                2 LITERAL "get"
                        3 SENSE 3
                2 LITERAL "get"
                        3 SENSE 4
        1 EQ_RELATION "eq_generalization"
                2 TARGET_ILI
                        3 WORDNET_OFFSET 69344
                2 TARGET_ILI
                        3 WORDNET_OFFSET 69756
```

So in this example, the Composite ILI-record represents a generalization between two specific synsets, which are senses of the verb “get”.

Finally, we give an example of ILI import records for computer terminology that has been added. It has the same general structure of an ADD_ON record but no equivalence relations:

```
0 ILI_RECORD
       1 PART_OF_SPEECH "n"
       1 ADD_ON_ID 8001
                1 GLOSS "COMPUTER_TERMINOLOGY Redefining in a child class a method or
                function member defined in a parent class."
       1 VARIANTS
              2 LITERAL "overriding"
                     3 SENSE 1
```

2.5.3. Import format for Top-Concepts and Domains

The top-ontology, which will be explained below, has internal structure and is linked to the ILI as well. The import records therefore consist of:

- variants (only one)
- gloss
- internal links
- links to the ILI

The internal links are limited to SUPER_TOP_CONCEPT, which stands for hyponymy, isa or superordinate, and OPPOSITE_TOP_CONCEPT to indicate explicit disjointness of classes:

```
0 TOP_CONCEPT
   1 VARIANTS
       2 LITERAL "1stOrderEntity"
   1 GLOSS "Any concrete entity (publicly) perceivable by the senses and located at
           any point in time, in a three-dimensional space."
   1 INTERNAL_LINKS
      2 SUPER_TOP_CONCEPT "Top"
      2 OPPOSITE_TOP_CONCEPT "2ndOrderEntity"
      2 OPPOSITE_TOP_CONCEPT "3rdOrderEntity"
   1 ILI_LINKS
      2 TARGET_ILI
         3 PART_OF_SPEECH "n"
         3 FILE_OFFSET 1958400

0 TOP_CONCEPT
   1 VARIANTS
      2 LITERAL "2ndOrderEntity"
   1 GLOSS "Any Static Situation (property, relation) or Dynamic Situation, which
            cannot be grasped, heart, seen, felt as an independent physical thing.
            They can be located in time and occur or take place rather than exist;
            e.g. continue, occur, apply"
   1 INTERNAL_LINKS
      2 SUPER_TOP_CONCEPT "Top"
      2 OPPOSITE_TOP_CONCEPT "1stOrderEntity"
      2 OPPOSITE_TOP_CONCEPT "3rdOrderEntity"
```

Domain labels can be imported in the same way, as hierarchical structures related to specific ILI-records. Except for the fact that the first line should say 0 DOMAIN and the concept internal relation is SUBDOMAIN:

```
0 @1@ DOMAIN
  1 GLOSS "hardware, software and elements from related scientific disciplines"
  1 VARIANTS
    2 LITERAL "computing"
  1 INTERNAL_LINKS
    2 SUB_DOMAIN "World-Wide Web"
    2 SUB_DOMAIN "networking"
    2 SUB_DOMAIN "storage"
    2 SUB_DOMAIN "programming"
    2 SUB_DOMAIN "operating system"
    2 SUB_DOMAIN "hardware"
  1 ILI_LINKS
    2 TARGET_ILI
      3 PART_OF_SPEECH "n"
      3 WORDNET_OFFSET 4339459
    2 TARGET_ILI
      3 PART_OF_SPEECH "n"
      3 WORDNET_OFFSET 2393633
```

3. Methodology

3.1. Expand/Merge approach

The EuroWordNet database was built (as much as possible) from available existing resources and databases with semantic information developed in various projects. This was not only more cost-effective given the limited time and budget of the project, but also made it possible to combine information from independently created wordnets.

In general, the wordnets were built in two major cycles as indicated by I and II in Figure 9 below. Each cycle consisted of a building phase and a comparison phase:

1. Building a wordnet fragment
1.1. Specification of an initial vocabulary
   1.2. Encoding of the language-internal relations
   1.3. Encoding of the equivalence relations
2. Comparing the wordnet fragments
   2.1. Loading of the wordnets in the EuroWordNet database
   2.2. Comparing and restructuring the fragments
   2.3. Measuring the overlap across the fragments

The building of a fragment was done using local tools and databases that are tailored to the specific nature and possibilities of the available resources. The available resources differ considerably in quality and explicitness of the data. Whereas some sites had the availability of partially structured networks between word senses, others started from genus words extracted from definitions that still had to be disambiguated in meaning.

After the specification of a fragment of the vocabulary, where each site used similar criteria (there may again be differences due to the different starting points), globally, two approaches have been followed for encoding the semantic relations:

> **Merge Model**: the selection is done in a local resource and the synsets and their language-internal relations are first developed separately, after which the equivalence relations to WordNet1.5 are generated.
>
> **Expand Model**: the selection is done in WordNet1.5 and the WordNet.1.5 synsets are translated (using bilingual dictionaries) into equivalent synsets in the other language. The wordnet relations are taken over and where necessary adapted to EuroWordNet. Possibly, monolingual resources are used to verify the wordnet relations imposed on non-English synsets.

The Merge Model, which was followed for most languages, results in a wordnet that is independent of WordNet1.5, possibly maintaining the language-specific properties. The Expand model, which was for example followed for Spanish and French, results in a wordnet that is very close to WordNet1.5 but which is also biased by it. What approach should be followed also depends on the quality of the available resources.

After the first production phase (steps Ia and Ib in Figure 9) the results have been converted to the EuroWordNet import format and loaded into the common database (step Ic). At that point various consistency checks have been carried out, both formally and conceptually. By using the specific options in the EuroWordNet database it is then possible to further inspect and compare the data, to restructure relations where necessary and to measure the overlap in the fragments developed at the separate sites.

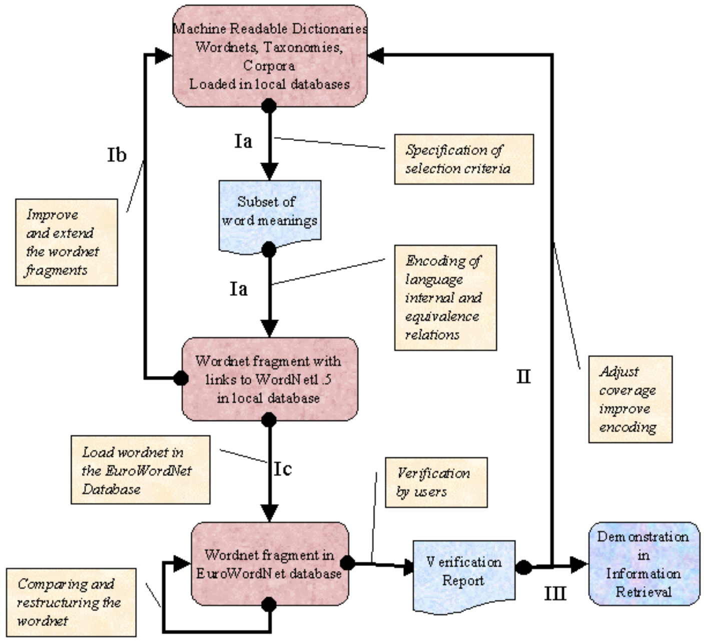


Figure 9: Global overview of steps in building EuroWordNet

After each cycle, there has been a verification phase. Feedback from the verification has been incorporated in the next building cycle. At the end of the project the results have been used in a (cross-language) information retrieval application (phase III).

The overall design of the EuroWordNet database made it possible to develop the individual language-specific wordnets relatively independently while guaranteeing a minimal level of compatibility. Nevertheless, some specific measures have been taken to enlarge the compatibility of the different resources:

1. The definition of a common set of so-called Base Concepts that is used as a starting point by all the sites to develop the cores of the wordnets. Base Concepts are meanings that play a major role in the wordnets.
2. The classification of the Base Concepts in terms of a Top Ontology.
3. The exchange of problems and possible solutions for encoding the relations for the Base Concepts.

Below we will give a further specification of the procedure of selecting the Base Concepts and the Top Ontology that has been used to classify them. In Vossen et al. (1998) a description is given of the kind of problems that have been encountered encoding relations in EuroWordNet and of the solutions that have been adopted.

3.2. Base Concepts

The main characteristic of the Base Concepts is their importance in the wordnets. According to our pragmatic point of view, a concept is important if it is widely used, either directly or as a reference for other widely used concepts. Importance is thus reflected in the ability of a concept to function as an anchor to attach other concepts. This anchoring capability has been defined in terms of two operational criteria that can be automatically applied to the available resources:

- the number of relations (general or limited to hyponymy).
- high position of the concept in a hierarchy (in WN1.5 or in any local taxonomy)

The notion of Base Concepts should thus not be confused with Basic-Level Concepts as defined by Rosch (1977). According to Rosch, the Basic Level is the level at which two conflicting principles of classification are in balance: 1) to predict features for as many instances as possible, 2) to predict as many features as possible. Typically, this balance occurs at an average level of specificity (where the level can vary due to interest and experience). Base Concepts are technically defined as the concepts with most relations. This more strongly correlates with the first principle, and they are therefore in most cases more general than the Basic Level Concepts.

Because the selection of these concepts should not be biased by a particular language or resource, each site has carried out an independent selection in their language. These selections have been translated to the closest equivalents in WordNet1.5 and the translated selections have been compared.

We first made a comparison between the Base Concepts (BCs) selected in the English, Dutch, Italian and Spanish wordnets. This set has been verified later by taking similar selections in the German, Estonian and Czech wordnets (for the French wordnet no independent selection has been carried out).

Once each group had selected their local set of BCs and linked it to WN1.5 synsets, we have computed the different intersections (pairs, triples, etc.) of the local BCs. In the ideal case, the selected sets of concepts coincide. The intersection of the English (GB), Dutch (NL), Spanish (ES) and Italian (IT) translations was however only 30 BCs (24 noun synsets, 6 verb synsets). This total intersection is not a reliable set. Important concepts such as *animal*, *object*, *place*, *location* are not included. We therefore selected all concepts occurring in two sets: the intersection-pairs.

Table 8: Intersection-pairs of translations of English, Dutch, Spanish and Italian Base Concepts

	Nouns				Verbs			
	NL	**ES**	**IT**	**GB**	**NL**	**ES**	**IT**	**GB**
NL	1027	103	182	333	323	36	42	86
ES	103	523	45	284	36	128	18	43
IT	182	45	334	167	42	18	104	39
GB	333	284	167	1296	86	43	39	236

Merging these intersections resulted in a set of 871 WN1.5-synsets (694 nouns and 177 verbs) out of a total set of 2860 synsets. Inspection of the rejected cases resulted in an extension of the BC set with another 211 noun and 62 verb synsets. The total set of common BCs (CBCs), based on English, Dutch, Italian and Spanish, thus consisted of 1144 synsets, 905 nominal BCs and 239 verbal BCs.

This set of CBCs has been verified by the Base Concept selections extracted in a similar way in French (FR), German (DE), Estonian (EE) and Czech (CZ). Table 9 shows the complete intersection of the new selections and the selections made for Dutch, Spanish, Italian and English.

Table 9: Complete Intersections of Base Concept Selections

	Nouns	Verbs
Intersection GB, NL, IT, ES	24	6
Intersection FR, DE, EE, CZ	70	30
Intersection All	13	2

As before, the total intersection of BCs derived for the new languages (FR, DE, EE, CZ) is small (100 synsets). The total intersection by 8 languages is only 15 synsets. The union of the intersection pairs is a set of 877 synsets (619 nouns and 258 verbs), which is comparable with the union of intersection pairs for GB, NL, ES, and IT. The next two tables show how the new selections (EWN2) overlap with the first set of common BCs (EWN1).

Table 10: Overlap of EWN2 nouns and EWN1 nouns (905 CBCs)

NOUNS	*Local NBCs*	*Intersection with CNBC-ewn1 (905)*	*% of CNBC4-EWN1*	*% of Local BCs*	*NEW BCs (not in EWN1)*
FR	787	787	99,24%	100,00%	0
DE	460	202	25,47%	43,91%	258
CZ	726	271	34,17%	37,33%	455
EE	703	389	49,05%	55,33%	314
Union (selected by at least 1 side)	1727	811	102,27%	46,96%	916
Union of Intersection pairs (selected by at least 2 sides)	619	516	65,07%	83,36%	105
Intersection (selected by 4 sides)	70	70	8,83%	100,00%	

Table 11: Overlap of EWN2 verbs and EWN1 verbs (239 CBCs)

VERBS	*Local VBCs*	*Intersection with CVBC-ewn1 (239)*	*% of CNBC4-EWN1*	*% of Local BCs*	*New BCs (Not in EWN1)*
FR	225	225	94.14%	100.00%	0
DE	321	98	41.00%	30.53%	223
EE	459	145	60.67%	31.80%	314
CZ	260	71	29.71%	27.31%	189
Union (selected by at least 1 side)	872	233	97.49%	26.72%	639
Union of Intersection pairs (selected by at least 2 sides)	258	179	74.90%	69.38%	61
Intersection (selected by 4 sides)	30	30	12.55%	100.00%	

When we look at the individual selections, we see that the French selection fully overlaps with the CBCs in EWN1. This is due to the fact that they have directly translated the CBCs from EWN1 and did not make an independent selection. The other selections show an overlap between 34-54% for nouns and 27-30% for verbs. If we compare the union of the intersection pairs we see a much higher overlap: 83% for nouns and 69% for verbs. These synsets are thus selected for 4 or more languages. There appears to be a high overlap between the Base Concepts in EWN1 and EWN2.

There are 105 nouns and 61 verbs selected by at least 2 EWN2 sides that are not part of the set of common Base Concepts selected in EWN1. These have been added to the set of common Base Concepts, resulting in a final total of 1310 synsets: 1010 nominal and 300 verbal synsets. Note that this set does not represent the most minimal set of concepts. No attempts have been made to reduce the set by generalizing unbalanced selections (e.g. *dog* is selected but not *cat*), merging synomous concepts (e.g. *act* and *action*). The main idea of the selection has been to be complete rather than to be minimal.

Given this set of common Base Concepts, the local selections can be divided into:

- synsets that have been selected as CBC. This means that at least one other site considered this concept as basic.
- rejected, i.e. no other site has considered the concept as basic. The concept is not a CBC but it can still be part of the local BCs.
- missing, i.e. synsets selected by at least two other sites but not part of the local set

The result of this division for each group is given in the next table

Table 12: Selected and Rejected Base Concepts for each language

	Nouns				Verbs			
	Proposed	Selected	Rejected	Missing	Proposed	Selected	Rejected	Missing
NL	1027	429	598	265	323	126	197	51
ES	523	323	200	371	128	72	56	105
IT	334	239	95	455	104	63	41	114
GB	1296	594	702	100	236	132	104	45
FR	787	787	0	223	225	225	0	75
DE	460	261	199	794	321	139	182	161
EE	703	465	238	545	459	205	254	95
CZ	726	351	375	599	260	98	162	202

This table illustrates that, for instance in the case of the Dutch (NL) nouns, 429 out of 1027 candidates (local BCs) were selected (as being members of at least one other selection) and 598 were rejected. The fourth column indicates that 265 nominal senses of the common BCs were missing in the local Dutch selection.

Each group tried to represent the missing BCs as good as possible by the equivalent concepts in their language. The results of representing the common BCs in Spanish, Italian and Dutch is given below, where the BCs are measured in WordNet1.5 synsets.

Table 13: Number of Common Base Concepts represented in the local wordnets

	Local Synsets Related to CBCs	Eq_synonym Relations	Eq_near_ Synonym relations	CBCs Without Direct Equivalent
NL	992	725	269	97
ES	1012	1009	0	15
IT	878	759	191	9

The final column gives the BCs that could not directly be represented in the local wordnets. In total 105 CBCs could not been represented in all three wordnets, 12 of which not in two wordnets:

Table 14: Base Concept Gaps in at least two wordnets

body covering#1	Mental object#1; cognitive content#1; content#2
body substance#1	Natural object#1
social control#1	Place of business#1; business establishment#1
change of magnitude#1	Plant organ#1
contractile organ#1	Plant part#1
spatial property#1; spatiality#1	Psychological feature#1

The table clearly shows that the unrelated CBCs are in many cases multiwords in WordNet1.5 that either represent artificial word senses, or very technical word senses.

If there is no eq_synonym or eq_near_synonym for a CBC, it is still linked to the closest meaning in the local wordnet via a so-called complex equivalence relation, e.g.:

{ongelukkig#1}, Adjective (unhappy)
    EQ_STATE_OF unfortunate#1, unfortunate person#1, Noun
{onwel#1}, Adjective (sick)
    EQ_IS_CAUSED_BY cause to feel unwell#1, Verb
{bevatten#1}, Verb, (to contain)
    EQ_INVOLVED vessel#2, Noun

Just as a single meaning in the local wordnet may be related to several CBCs, it is also possible that a single CBC is related to several meanings in the local wordnets. Especially when it represents an intermediate level of classification, it makes sense to link the CBC both to a more general meaning in the local wordnet (with an eq_has_hyponym relation with the CBC) and to the more specific meanings that it classifies (with an eq_has_hyperonym relation the CBC). This is illustrated by the way in which the non-lexicalized BC “plant part” (0976849-n) is represented in the Spanish wordnet by linking hyponymic and holonymic Spanish synsets to it:

{cosa#1; objeto#1} Noun (inanimate object, physical object, object)
    EQ_HAS_HYPONYM plant part#1, Noun
{organo#5; organo vegetal#1}, Noun (plant organ)
    EQ_HAS_HYPERONYM plant part#1, Noun
{floar#1, planta#1} Noun, (plant life, flora, plant)
    EQ_HAS_MERONYM plant part#1, Noun

Via the complex equivalence relations we thus get a maximal coverage of all the CBCs by all the sites in terms of local representatives, even when there is no direct equivalence.

For building the wordnets, the meanings directly related to the CBCs are taken as the starting point in the local wordnet. These selections are then worked out according to the lexicalization patterns that are relevant to that particular language. It may turn out that some meanings related to a CBC are not important for the local wordnet. In that case, only the minimal relations are encoded (synonymy and hyponymy). It may also be the case that important meanings in the local wordnet are not part of the CBC-related set. In that case, they are given the same attention as the CBC-related meanings. The resulting core wordnet in each language will thus include the meanings related to the CBCs and any other meaning which is important for the wordnet.

Given the set of common BCs (1310), each site created their core wordnets independently using the following procedure (see Figure 10 for an overview):

1. extend the set of Local BCs with equivalent representatives for the missing BCs.
2. create synsets for the Local BCs and the common BC (CBC) representatives.
3. encode the hyperonyms for the Local BCs and the CBC representatives (as far as they are not yet part of the selection).
4. encode the first level of hyponyms below the Local BCs and the CBC representatives
5. encode synsets related to the Local BCs and CBC representatives by non-hyponymy relations
6. encode sub-hyponyms of the Local BCs and CBC representatives

Figure 10 gives an overview of the different vocabulary fragments. Step 1 through 4 result in the core wordnets that are most important. We have focussed the manual work on the core wordnets. Extensions from the core make it possible to apply different (semi-)automatic methodologies for building and to include language specific lexicalization patterns. As indicated in the general scheme, the intermediate results have been compared. The results have been used to adjust the building strategies.

The documents that accompany each wordnet further describe the building and selection of the different vocabularies and how they are compared. Each site has been free to add other concepts to the core wordnets, suiting their local approach and starting point. These additions could be:

- synsets related via non-hyponymy relations (such as meronymy, role/involvement, antonymy).
- synsets that are translatable to WordNet1.5 synsets.
- easily extractable from the lexical resources that are available.
- local Base Concepts, locally important concepts but still not part of the set of common Base Concepts.

For each of these synsets the following information has to be minimally specified:

- Hyperonym
- Synonyms (synset members)
- Equivalence relations to WordNet1.5, either directly or via a hyperonym

Optionally, any other relation could be added.

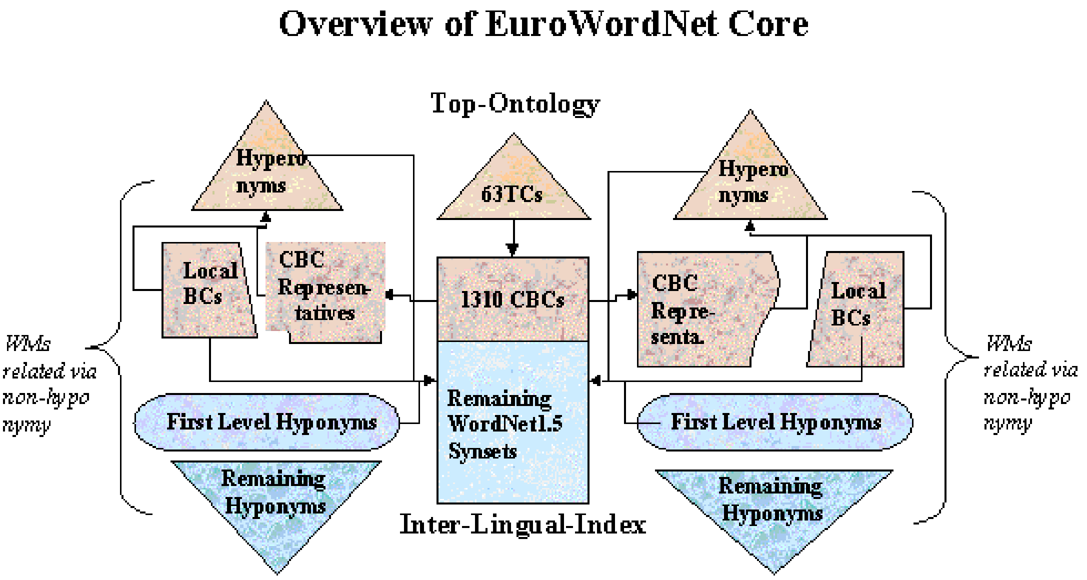


Figure 10: General outline of the vocabularies

Finally, as can be seen in Figure 10, the BCs have been clasified by the top-ontology of 63 semantic distinction. This ontology, which functions as a common framework for all the wordnets, will be described in the next section.

3.3. Top Ontology

To get to grips with the set of Base Concepts and to achieve consensus on the interpretation, we have constructed a top-ontology of basic semantic distinctions to classify them. As explained in the introduction, the language-specific modules (as autonomous systems of language-internal relations), are linked through the ILI, which gives further access to all language-independent knowledge, among which the Top Ontology of fundamental semantic distinctions. This language-independent information can be transferred via the ILI-records to all the language specific synsets that are linked to it. The common BCs, described above, are all specified in the form of ILI-records, which are thus linked to fundamental concepts in the local wordnets.

The purpose of the EuroWordNet Top Ontology can then be detailed as follows:

a) It enforces more uniformity and compatibility of the different wordnets. The classifications of the BCs in terms of the Top Ontology distinctions should apply to all the involved languages. In practice this means that all sites have verified the assignment of a Top Concept to an ILI-record for the synsets in their local wordnets that are linked to this ILI-record. For example, the features associated with the top-concept Object can only apply to the ILI-record *object*, when the features also apply to the Dutch and Italian concepts linked to this ILI-record as equivalences. In addition the distinction should also hold for all other Dutch and Italian concepts that could possibly inherit this property from the language-internal relations (e.g. all the (sub)hyponyms linked to “voorwerp” in the Dutch wordnet and all the (sub)hyponyms linked to “oggetto” in the Italian wordnet). Note that the language internal distribution of such a feature can still differ from wordnet to wordnet, as long as no false implications are derived.

b) Using the Top Concepts (TCs) we can divide the Base Concepts (BCs) into coherent clusters. This means that the building of the wordnets can take place from cluster to cluster so that similar concepts are dealt with adjacently. This is important to enable contrastive-analysis of the word meanings and it will stimulate a similar treatment. Furthermore, the clusters are used to monitor progress across the sites and to discuss problems and solutions per cluster.

c) The Top-Ontology provides users access and control of the database without having to understand the languages of the wordnets. It is possible to customize the database by assigning features to the top-concepts, irrespective of the language-specific structures.

d) Although the wordnets in EWN are seen as autonomous language-specific structures, it is in principle possible to extend the database with language-neutral ontologies, such as CYC, MikroKosmos, the Upper-Model, by linking them to the corresponding ILI-records. Such a linking will be facilitated by the top-concept ontology where similar concepts can be mapped directly.

From these purposes we can derive a few more specific principles for deciding on the relevant distinctions. As suggested before, the wordnets reflect language-specific dependencies between words. Likewise, the coding of the relations can be seen mainly as a linguistic operation, resulting in linguistically-motivated relations.[8] It is therefore important that the top-ontology incorporates semantic distinctions that play a role in linguistic approaches rather than purely cognitive or knowledge-engineering practices. We therefore have initially based the ontology on semantic classifications common in linguistic paradigms: Aktionsart models [Vendler 1967, Verkuyl 1972, Dowty 1979, Verkuyl 1989, Pustejovsky 1991, Levin 1993], entity-orders [Lyons 1977], Aristotle's Qualia-structure [Pustejovsky 1995]. Furthermore, we made use of ontological classifications developed in previous EC-projects, which had a similar basis and are well-known in the project consortium: Acquilex (BRA 3030, 7315), Sift (LE-62030, [Vossen and Bon 1996].[9]

In addition to these theoretically-motivated distinctions there is also a practical requirement that the ontology should be capable of reflecting the diversity of the set of common BCs, across the 8 languages. In this sense the classification of the common BCs in terms of the top-concepts should result in:

- homogeneous Base Concept Clusters
- average size of Base Concept Clusters

Homogeneity has been verified by checking the clustering of the BCs with their classification in WordNet1.5. In this senses the ontology has also been adapted to fit the top-levels of WordNet1.5. Obviously, the clustering also has been verified with the other language-specific wordnets. The criterion of cluster-size implies that we should not get extremely large or small clusters. In the former case the ontology should be further differentiated, in the latter case distinctions have to be removed and the BCs have to be linked to a higher level. Finally, we can mention as important characteristics:

- the semantic distinctions should apply to both nouns, verbs and adjectives, because these can be related in the language-specific wordnets via a xpos_synonymy relation, and the ILI-records can be related to any part-of-speech.
- the top-concepts are hierarchically ordered by means of a subsumption relation but there can only be one super-type linked to each top-concept: multiple inheritance between top-concepts is not allowed.
- in addition to the subsumption relation, top-concepts can have an opposition-relation to indicate that certain distinctions are disjunct, whereas others may overlap.
- there may be multiple relations from ILI-records to top-concepts. This means that the BCs can be cross-classified in terms of multiple top-concepts (as long as these have no opposition-relation between them): i.e. multiple inheritance from Top-Concept to Base Concept is allowed.

It is important to realize that the Top Concepts (TCs) are more like semantic features than common conceptual classes. We typically find TCs for Living and for Part but we do not find a TC Bodypart, even though this may be more appealing to a non-expert. BCs representing *body parts* are now cross-classified by two feature-like TCs Living and Part. The reason for this is that the diversity of the BCs would require many cross-classifiying concepts where Living and Part are combined with many other TCs. These combined classes result in a much more complex system, which is not very flexible and difficult to maintain or adapt. Furthermore, it turned out that the BCs typically abstract from particular

[8] Relations hold between lexicalized units (words and phrases) of a language, and not, as is often the case in language-neutral ontologies, just for the sake of creating a better ordering of hierarchies. The wordnets should therefore not contain levels or synsets for concepts which are not considered to be natural expressions in a language; this to the contrary of the common practice in WordNet1.5. As linguistic-structures the wordnets can provide valuable information on the expressiveness of languages, as conceptual-structures this is not guaranteed.

[9] In a later stage the EWN ontology has been compared with language-neutral ontologies such as CYC, Upper-Model, MikroKosmos. This took place in the framework of the Eagles-project and in collaboration with the ANSI ADHOC Group on Ontology Standards.

features but these abstractions do not show any redundancy: i.e. it is not the case that all things that are Living also always share other features.

An explanation for the diversity of the BCs is the way in which they have been selected. To be useful as a classifier or category for many concepts (one of the major criteria for selection) a concept must capture a particular generalization but abstract from (many) other properties. Likewise we find many classifying meanings which express only one or two TC-features but no others. In this respect the BCs typically abstract one or two levels from the cognitive Basic-Level as defined by [Rosch 1977]. So we more likely find BCs such as *furniture* and *vehicle* than *chair*, *table* and *car*.

The ontology is the result of 4 cycles of updating where each proposal has been verified by the different sites. The ontology now consists of 63 higher-level concepts, excluding the top. Following [Lyons 1977] we distinguish at the first level 3 types of entities:

1stOrderEntity
Any concrete entity (publicly) perceivable by the senses and located at any point in time, in a three-dimensional space, e.g.: vehicle, animal, substance, object.

2ndOrderEntity
Any Static Situation (property, relation) or Dynamic Situation, which cannot be grasped, heard, seen, felt as an independent physical thing. They can be located in time and occur or take place rather than exist, e.g.: happen, be, have, begin, end, cause, result, continue, occur..

3rdOrderEntity
Any unobservable proposition which exists independently of time and space. They can be true or false rather than real. They can be asserted or denied, remembered or forgotten, e.g.: idea, thought, information, theory, plan.

According to Lyons, 1stOrderEntities are publicly observable individual persons, animals and more or less discrete physical objects and physical substances. They can be located at any point in time and in, what is at least psychologically, a three-dimensional space. The 2ndOrderEntities are events, processes, states-of-affairs or situations which can be located in time. Whereas 1stOrderEntities **exist** in time and space 2ndOrderEntities **occur** or **take place**, rather than exist. The 3rdOrderEntities are propositions, such as ideas, thoughts, theories, hypotheses, that exist outside space and time and which are unobservable. They function as objects of propositional attitudes, and they cannot be said to occur or be located either in space or time. Furthermore, they can be predicated as true or false rather than real, they can be asserted or denied, remembered or forgotten, they may be reasons but not causes.

The following tests are used to distinguish between 1st and 2nd order entities:

a The same person was here again today
b The same thing happened/occurred again today

The reference of 'the same person' is constrained by the assumption of spatio-temporal continuity and by the further assumption that the same person cannot be in two different places at the same time. The same *type* of event can occur in several different places, not only at different times but also at the same time. However, the same event cannot reoccur at all; it is for allways bound by the time and location of its occurrence. Third-order entities cannot occur, have no temporal duration and therefore fail on both tests:

*? The idea, fact, expectation, etc.... was here/occurred/ took place

A positive test for a 3rdOrderEntity is based on the properties that can be predicated:

ok The idea, fact, expectation, etc.. is true, is denied, forgotten

The first division of the ontology is disjoint: BCs cannot be classified as combinations of these TCs. This distinction cuts across the different parts of speech in that:

- 1stOrderEntities are always (concrete) nouns.
- 2ndOrderEntities can be nouns, verbs or adjectives.
- 3rdOrderEntities are always (abstract) nouns.

The actual distribution of the BCs over the different parts of speech is shown in the next table:

Table 15: Total Set of classified Base Concepts

	Nouns	Verbs	Total
1stOrderEntities	491		491
2ndOrderEntities	272	263	535
3rdOrderEntities	33		33
Total	796	263	1059

The figures given here and below cover the Base Concepts before the extension based on the French, German, Czech and Estonian selections. Note also that a BC may originally be a noun or verb in WordNet1.5 but may be associated with any part-of-speech in one of the local wordnets. The 1stOrderEntities and 2ndOrderEntities are then further subdivided according to the following hierarchy, where the superscripts indicate the number of BCs that are directly classified by the TC:

Top^{0}

1stOrderEntity^{1}

- **Origin^{0}**
  - Natural^{21}
    - Living^{30}
      - Plant^{18}
      - Human^{106}
      - Creature^{2}
      - Animal^{23}
  - Artifact^{144}
- **Form^{0}**
  - Substance^{32}
    - Solid^{63}
    - Liquid^{13}
    - Gas^{1}
  - Object^{162}
- **Composition^{0}**
  - Part^{86}
  - Group^{63}
- **Function^{55}**
  - Vehicle^{8}
  - Representation^{12}
    - MoneyRepresentation^{10}
    - LanguageRepresentation^{34}
    - ImageRepresentation^{9}
  - Software^{4}
  - Place^{45}
  - Occupation^{23}
  - Instrument^{18}
  - Garment^{3}
  - Furniture^{6}
  - Covering^{8}
  - Container^{12}
  - Comestible^{32}
  - Building^{13}

2ndOrderEntity^{0}

- **SituationType^{6}**
  - Dynamic^{134}
    - BoundedEvent^{183}
    - UnboundedEvent^{48}
  - Static^{28}
    - Property^{61}
    - Relation^{38}
- **SituationComponent^{0}**
  - Cause^{67}
    - Agentive^{170}
    - Phenomenal^{17}
    - Stimulating^{25}
  - Communication^{50}
  - Condition^{62}
  - Existence^{27}
  - Experience^{43}
  - Location^{76}
  - Manner^{21}
  - Mental^{90}
  - Modal^{10}
  - Physical^{140}
  - Possession^{23}
  - Purpose^{137}
  - Quantity^{39}
  - Social^{102}
  - Time^{24}
  - Usage^{8}

3rdOrderEntity^{33}

Figure 11: The EuroWordNet Top-Ontology

Since the number of 3rdOrderEntities among the BCs was limited compared to the 1stOrder and 2ndOrder Entities we have not further subdivided them. The following BCs have been classified as 3rdOrderEntities:

> **Base Concepts classified as 3rdOrderEntities:**
> theory; idea; structure; evidence; procedure; doctrine; policy; data point; content; plan of action; concept; plan; communication; knowledge base; cognitive content; know-how; category; information; abstract; info;

The subdivisions of the 1stOrderEntities and 2ndOrderEntities are further discussed in the next sections.

3.3.1. Classification of 1st-Order-Entities

The 1stOrderEntities are distinguished in terms of four main ways of conceptualizing or classifying a concrete entity:

a) Origin: the way in which an entity has come about.
b) Form: as an a-morf substance or as an object with a fixed shape, hence the subdivisions Substance and Object.
c) Composition: as a group of self-contained wholes or as a part of such a whole, hence the subdivisions Part and Group.
d) Function: the typical activity or action that is associated with an entity.

These classes are comparable with Aristotle's Qualia roles as described in Pustejovsky's Generative lexicon, (the Agentive role, Formal role, Constitutional role and Telic Role respectively: [Pustejovsky 1995] but are also based on our empirical findings to classify the BCs. BCs can be classified in terms of any combination of these four roles. As such the top-concepts function more as features than as ontological classes. Such a systematic cross-classification was necessary because the BCs represented such diverse combinations (e.g. it was not possible to limit Function or Living only to Object).

The main-classes are then further subdivided, where the subdivisions for Form and Composition are obvious given the above definition, except that Substance itself is further subdivided into Solid, Liquid and Gas. In the case of Function the subdivisions are based only on the frequency of BCs having such a function or role. In principle the number of roles is infinite but the above roles appear to occur more frequently in the set of common Base Concepts.

Finally, a more fine-grained subdivision has been made for Origin, first into Natural and Artifact. The category Natural covers both inanimate objects and substances, such as *stones*, *sand*, *water*, and all living things, among which *animals*, *plants* and *humans*. The latter are stored at a deeper level below Living. The intermediate level Living is necessary to create a separate cluster for natural objects and substances, which consist of Living material (e.g. *skin*, *cell*) but are not considered as *animate beings*. Non-living and Natural objects and substances, such as natural products like *milk*, *seeds*, *fruit*, are classified directly below Natural.

As suggested, each BC that is a 1stOrderEntity is classified in terms of these main classes. However, whereas the main-classes are intended for cross-classifications, most of the subdivisions are disjoint classes: a concept cannot be an Object and a Substance, or both Natural and Artifact. This means that within a main-class only one subdivision can be assigned. Consequently, each BC that is a 1stOrderEntity has at least one up to four classifications:

fruit:	Comestible (Function) Object (Form) Part (Composition) Plant (Natural, Origin)
skin:	Covering (Covering) Solid (Form) Part (Constituency) Living (Natural, Origin)
life 1:	Group (Composition) Living (Natural, Origin)

cell:	Part (Composition)
	Living (Natural, Origin)
reproductive structure 1	Living (Natural, Origin)

The next Figure give a schematic overview, how clusters of BCs (both 1stOrder and 2ndOrderEntites) are classified by combinations of TCs:

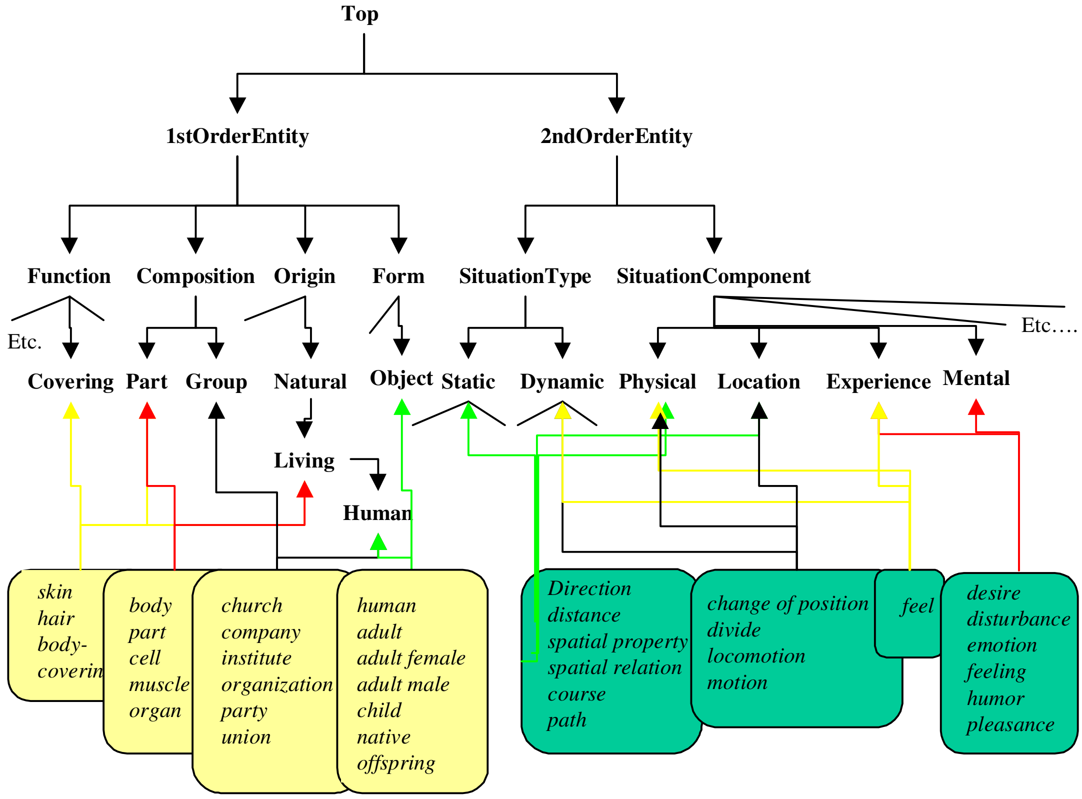


Figure 12: Lattice structure of the EuroWordNet top-ontology

The more classifications apply, the more informative the concept is. If a BC is classified by e.g. only one main-class it means that it can refer to things that vary in properties with respect to the other classes. This typically applies to words which we call Functionals and which occur relatively often as BCs. Functionals are words that can only be characterized in terms of some major activity-involvement and can vary with respect to their Form, Constituency, or Origin. Examples of Functionals are: *threat*, *belongings*, *product*, *cause*, *garbage*, which can refer to persons, animals, substances, objects, instruments, parts, groups, anything as long as it satisfies the described role. These nouns thus have an open denotation (although stereotypical constraints may hold) and fully rely on this role relation.[10] Other classes below Function, e.g. Building, Vehicle are also linked to Artifact and therefore specified for Origin. Most of these are Objects, some are also specified for Group:

arms:	Instrument (Function)
	Group (Composition)
	Object (Form)
	Artifact (Origin)

Finally, with respect to Composition it needs to be said that only concepts that essentially depend on some other concept, are classified as either Part or Group. It is not the case that all *persons* will be classified as Parts because they may be part of *group*. *Group*, on the other hand, typically depends on the elements as part of its meaning.

[10] This role relation may be expressed in the language-internal wordnet by means of a specific role-relation with a lexicalized verb or noun denoting the event.

Table 16: Definitions for first order top concepts

1stOrder Top Concept	*Gloss*
Origin	Considering the way concrete entities are created or come into existence.
Function	Considering the purpose, role or main activity of a concrete entity. Typically it can be used for nouns that can refer to any substance, object which is involved in a certain way in some event or process; e.g. remains, product, threat.
Form	Considering the shape of concrete entities, fixed as an object or a-morf as a substance
Composition	Considering the composition of concrete entities in terms of parts, groups and larger constructs
Part	Any concrete entity which is contained in an object, substance or a group; head, juice, nose, limb, blood, finger, wheel, brick, door
Group	Any concrete entity consisting of multiple discrete objects (either homogeneous or heterogeneous sets), typically people, animals, vehicles; e.g. traffic, people, army, herd, fleet
Substance	all stuff without boundary or fixed shape, considered from a conceptual point of view not from a linguistic point of view; e.g. mass, material, water, sand, air. Opposed to Object.
Object	Any conceptually-countable concrete entity with an outer limit; e.g. book, car, person, brick. Opposed to Substance.
Vehicle	; e.g. car, ship, boat
Software	; e.g. computer programs and databases
Representation	Any concrete entity used for conveying a message; e.g. traffic sign, word, money.
Place	Concrete entities functioning as the location for something else; e.g. place, spot, centre, North, South
Occupation	; e.g. doctor, researcher, journalist, manager
Instrument	; e.g. tool, machine, weapon
Garment	; e.g. jacket, trousers, shawl
Furniture	; e.g. table, chair, lamp
Covering	; skin, cloth, shield,
Container	; e.g. bag, tube, box
Comestible	food & drinks, including substances, liquids and objects.
Building	; e.g. house, hotel, church, office
Plant	; e.g. plant, rice; Opposed to Animal, Human, Creature.
Human	; e.g. person, someone
Creature	Imaginary creatures; e.g. god, Faust, E.T.; Opposed to Animal, Human, Plant
Animal	; e.g. animal, dog; Opposed to Plant, Human, Creature.
Living	Anything living and dying including objects, organic parts or tissue, bodily fluids; e.g. cells; skin; hair, organism, organs.
Natural	Anything produced by nature and physical forces as artifact; Opposed to Artifact.
Artifact	Anything manufactured by people as natural; Opposed to Natural.
MoneyRepresentation	Physical Representations of value, or money; e.g. share, coin
LanguageRepresentation	Physical Representations conveyed in language (e.g. spoken, written or sign language); e.g. text, word, utterance, sentence, poem
ImageRepresentation	Physical Representations conveyed in a visual medium; e.g. sign language, traffic sign, light signal
Solid	Substance which can fall, does not feel wet and you cannot inhale it; e.g. stone, dust, plastic, ice, metal; Opposed to Liquid, Gas
Liquid	Substance that can fall, feels wet and can flow on the ground; e.g. water, soup, rain; Opposed to Gas, Solid.
Gas	Substance that cannot fall, you can inhale it and it floats above the ground; e.g. air, ozon; Opposed to Liquid, Solid.

3.3.2. The classification of 2ndOrderEntities

As explained above, 2ndOrderEntities can be referred to using nouns and verbs (and also adjectives or adverbs) denoting static or dynamic Situations, such as *birth*, *live*, *life*, *love*, *die* and *death*. All 2ndOrderEntities are classified using two different classification schemes, which represent the first division below 2ndOrderEntity:

- the SituationType: the event-structure in terms of which a situation can be characterized as a conceptual unit over time;
- the SituationComponent: the most salient semantic component(s) that characterize(s) a situation;

The SituationType reflects the way in which a situation can be quantified and distributed over time, and the dynamicity that is involved. It thus represents a basic classification in terms of the event-structure (in the formal tradition) or the predicate-inherent Aktionsart properties of nouns and verbs. Examples of SituationTypes are Static, Dynamic. The SituationComponents represent a more conceptual classification, resulting in intuitively coherent clusters of word meanings. The SituationComponents reflect the most salient semantic components that apply to our selection of Base Concepts. Examples of SituationComponents are: Location, Existence, Cause.

Typically, SituationType represents disjoint features that cannot be combined, whereas it is possible to assign any range or combination of SituationComponents to a word meaning. Each 2ndOrder meaning can thus be classified in terms of an obligatory but unique SituationType and any number of SituationComponents.

3.3.2.1. SituationTypes

Following a traditional Aktionsart classification [Vendler 1967, Verkuyl 1972, Dowty 1979, Verkuyl 1989], SituationType is first subdivided into Static and Dynamic, depending on the dynamicity of the Situation:

Dynamic

Situations implying either a specific transition from one state to another (Bounded in time) or a continuous transition perceived as an ongoing temporally unbounded process; e.g. event, act, action, become, happen, take place, process, habit, change, activity. Opposed to Static.

Static

Situations (properties, relations and states) in which there is no transition from one eventuality or situation to another, i.e. they are non-dynamic; e.g. state, property, be. Opposed to Dynamic.

In general words, Static Situations do not involve any change, Dynamic Situations involve some specific change or a continuous changing. The traditional test for making dynamicity explicit is to combine the noun or verb with a manner phrase that specifies the inherent properties of the Situation:

a. ?he sits quickly.
b. he sat down quickly; a quick, wild meeting

The static verb *to sit* cannot be combined with quickly, but the dynamic verb *to sit down* and dynamic noun *meeting* can. Different aspectual modifications, such as (im)perfective, progressive, depend on this qualification.

Static Situations are further subdivided into Properties, such as *length*, *size,* which apply to single concrete entities or abstract situations, and Relations, such as *distance*, *space*, which only exist relative to and in between several entities (of the same order):

Property

Static Situation which applies to a single concrete entity or abstract Situation; e.g. colour, speed, age, length, size, shape, weight.

Relation

Static Situation which applies to a pair of concrete entities or abstract Situations, and which cannot exist by itself without either one of the involved entities; e.g. relation, kinship, distance, space.

Dynamic Situations are subdivided into events which express a specific transition and are bounded in time (BoundedEvent), and processes which are unbounded in time (UnboundedEvent) and do not imply a specific transition from one situation to another (although there can be many intermediate transitions):

BoundedEvent
: Dynamic Situations in which a specific transition from one Situation to another is implied; Bounded in time and directed to a result; e.g. to do, to cause to change, to make, to create.

UnboundedEvent
: Dynamic Situations occurring during a period of time and composed of a sequence of (micro-)changes of state, which are not perceived as relevant for characterizing the Situation as a whole; e.g. grow, continuous changing, move around, live, breath, activity, hobby, sport, education, work, performance, fight, love, caring, management.

We typically see that many verbs and nouns are under-classified for boundedness and sometimes even for dynamicity. This means that they can get a more specific interpretation in terms of a bounded change or an unbounded process when they are put in a particular context. A verb such as “walk” names a bounded event when it is combined with a destination phrase, as in (a), but it is unbounded when it is combined with a location phrase as in (b):

a) He walked to the station (?for hours) (in 2 hours)
b) He walked in the park (for hours) (?in 2 hours)

The boundedness is made more explicit using duration phrases that imply the natural termination point of the change (*in 2 hours*) or explicitly do not (*for hours*).

3.3.2.2 SituationComponents

The SituationComponents divide the Base-Concepts in conceptually coherent clusters. The set of distinctions is therefore based on the diversity of the set of common Base-Concepts that has been defined. The following main components have been distinguished (where each component is followed by a formal definition and a short explanation):

Usage
: Situations in which something (an instrument, substance, time, effort, force, money) is or can be used; e.g. to use, to spent, to represent, to mean, to be about, to operate, to fly, drive, run, eat, drink, consume.

Usage stands for Situations in which either a resource or an instrument is used or activated for some purpose. This covers both consumptive usage (the use time, effort, food, fuel) and instrumental operation (as in *to operate a vehicle*, *to run a program).* So far it has been restricted to Dynamic Situations only. It typically combines with Purpose, Agentive and Cause because we often deliberately use things to cause to some effect for some purpose.

Time
: Situations in which duration or time plays a significant role; Static yesterday, day, pass, long, period, Dynamic e.g. begin, end, last, continue.

Time is only applied to BCs that strongly imply temporal aspects. This includes general BCs that only imply some temporal aspect and specific BCs that also denote some specific Situation. Typical ‘aspectual’ BCs, such as *begin*, *end*, only express to the phase of situations but abstract from the actual Situation. Most of these also imply dynamicity. More specific BCs, such as to attack, *to depart, to arrive*, combine other SituationComponents but also imply some phase. Finally, all BCs that denote time points and periods, such as *time*, *day*, *hour*, *moment*, are all clustered below Time and Static.

Social
: Situations related to society and social interaction of people: Static e.g. employment, poor, rich, Dynamic e.g. work, management, recreation, religion, science.

Social refers to our inter-human activities and situations in society. There are many Social activities (UnboundedEvent) which correlate with many different Social Interests or Purposes. These are not further differentiated in terms of TCs but using the Domain labels (Management, Science, Religion, Health Care, War, Recreation, Sports). In addition there are Static Social states such as *poverty*, *employment*.

Quantity

Situations involving quantity and measure; Static e.g. weight, heaviness, lightness; changes of the quantity of first order entities; Dynamic e.g. to lessen, increase, decrease.

Dynamic BCs clustered below Quantity typically denote increase or decrease of amounts of entities. Static Quantity BCs denote all kinds of measurements.

Purpose

Situations which are intended to have some effect.

Purpose is an abstract component reflecting the intentionality of acts and activities. This concept can only be applied to Dynamic Situations and it strongly correlates with Agentive and Cause, clustering mainly human acts and activities. SituationComponents such as Usage, Social and Communication often (but not always) combine with Purpose.

Possession

Situations involving possession; Static e.g. have, possess, possession, contain, consist of, own; Dynamic changes in possession, often to be combined which changes in location as well; e.g. sell, buy, give, donate, steal, take, receive, send.

Possession covers ownership and changes of ownership, but not physical location or meronymy or abstract possession of properties. The fact that transfer of Possession often implies physical motion or static location will be indicated by cross-classifying BCs for Possession, Location, and Static or Dynamic, respectively.

Physical

Situations involving perceptual and measurable properties of first order entities; either Static e.g. health, a colour, a shape, a smell; or Dynamic changes and perceptions of the physical properties of first order entities; e.g. redden, thicken, widen, enlarge, crush, form, shape, fold, wrap, thicken, to see, hear, notice, smell. Opposed to Mental.

Physical typically clusters Dynamic physical Changes, in which a Physical Property is altered, and Static Physical Properties. In all these cases a particular physical property is incorporated which, in many cases, can be made explicit by means of a causative relation (*to become red*) or a synonymy relation (*health and healthy*) with an adjective in the local wordnets. Another cluster is formed by Physical Experiences (see Experience).

Modal

Situations (only Static) involving the possibility or likelihood of other situations as actual situations; e.g. abilities, power, force, strength.

Modal Situations are always Static. Most Modal BCs denote some ability or necessary property needed to perform some act or activity.

Mental

Situations experienced in mind, including a concept, idea or the interpretation or message conveyed by a symbol or performance (meaning, denotation, content, topic, story, message, interpretation) and emotional and attitudinal situations; a mental state is changed; e.g. invent, remember, learn, think, consider. Opposed to Physical.

Mental Situations can be differentiated into Experiences (see Experience) and in Dynamic Mental events possibly involving an Agent. The latter cluster cognitive actions and activities such as *to think, to calculate, to remember*, *to decide*.

Manner

Situations in which way or manner plays a role. This may be Manner incorporated in a dynamic situation, e.g. ways of movement such as walk, swim, fly, or the static Property itself: e.g. manner, sloppy, strongly, way.

Manner as a SituationComponent applies to many specific BCs that denote a specific way or manner in which a Dynamic event takes place. Typical examples are ways of movement. General BCs that only refer to Manner as such and not to some specific Situation are Static nouns such as *manner*, *way*, *style*.

Location

Situations involving spatial relations; static e.g. level, distance, separation, course, track, way, path; something changes location, irrespective of the causation of the change; e.g. move, put, fall, drop, drag, glide, fill, pour, empty, take out, enter.

Location is typically incorporated in Dynamic BCs denoting *movements*. When combined with Static it clusters nouns that refer to Location Relations, such as *distance, level, path, space*. A Location Relation holds between several entities and cannot be seen as a property of single entity. This makes it different from Place, which applies to a 1stOrderEntity that functions as the location for an event or some other 1stOrderEntity.

Experience

Situations that involve an experiencer: either mental or perceptual through the senses.

Situations with the TC Experience involve the mental or perceptual processing of some stimulus. In this respect there must be an experiencer implied, although it is not necessarily expressed as one of the arguments of a verb (it could be incorporated in the meaning). Typical Experience BCs are: *to experience, to sense, to feel, pain, to notice*. Experiences can be differentiated by combining it with Physical or Mental. Physical Experiences are external stimuli processed by the senses: *to see, to hear*. Mental Experiences are internal only existing in our minds: *desire, pleasance, humor, faith, motivation*. There are many examples of BCs that cannot be differentiated between these, e.g. *pain* that can be both Physical and Mental. Another interesting aspect of Experiences is that there is unclarity about the dynamicity. It is not clear whether a *feeling* or *emotion* is static or dynamic. In this respect Experience BCs are often classified as SituationType, which is undifferentiated for dynamicity.

Existence

Situations involving the existence of objects and substances; Static states of existence e.g. exist, be, be alive, life, live, death; Dynamic changes in existence; e.g. kill, produce, make, create, destroy, die, birth.

Dynamic Existence Situations typically refer to the coming about, the dying or destruction of both natural and artifact entities. This includes artificial production or creation, such as *to make, to produce, to create, to invent*, and natural *birth*. Static Existence is a small cluster of nouns that refer to existence or non-existence.

Condition

Situations involving an evaluative state of something: Static, e.g. health, disease, success or Dynamic e.g. worsen, improve.

Condition is an evaluative notion that can be either positive or negative. It can be combined with Dynamic changes (Social, Physical or Mental) or Static Situations which are considered as positive or negative (again Social, Physical or Mental).

Communication

Situations involving communication, either Static, e.g. be_about or Dynamic (Bounded and Unbounded); e.g. speak, tell, listen, command, order, ask, state, statement, conversation, call.

Communication verbs and nouns are often speech-acts (bounded events) or denote more global communicative activities (unbounded events) but there are also a few Static Communication BCs. The Static Communication BCs (e.g. *to be about*) express meaning relations between PhysicalRepresentations (such as written language) and the propositional content (3rdOrderEntities). The Dynamic BCs below the TC Communication form a complex cluster of related concepts. They can represent various aspects of Communication which correlate with the different ways in which the communication is brought about, or different phases of the communication. Some Communication BCs refer to causation of communication effects, such as *to explain, to show, to demonstrate*, but not necessarily to the precise medium (graphical, verbal, body expression). These BCs combine with the TCs Cause and Mental. Other BCs refer to the creation of a meaningful Representation, *to write, to draw, to say*, but they do not necessarily imply a communicative effect or the perception and interpretation of the Representation. They typically combine with Existence, Agentive, and Purpose. Yet other BCs refer to the perceptual and mental processing of communicative events, *to read, to listen* and thus combine with Mental.

Cause

Situations involving causation of Situations (both Static and Dynamic); result, effect, cause, prevent.

Causation is always combined with Dynamic and it can take various forms. It can either be related to a controlling agent which intentionally tries to achieve some change (Agentive), or it can be related to some natural force or circumstance (Phenomenal). Another differentiation is into the kind of effect as a perceptive or mental Experience, which makes the cause Stimulating. The different ways of causation have been subdivided in terms of an extra level of TCs:

Agentive

Situations in which a controlling agent causes a dynamic change; e.g. to kill, to do; to act. Opposed to other causes such as Stimuli, Forces, Chance, Phenomena.

Stimulating

Situations in which something elicits or arouses a perception or provides the motivation for some event, e.g. sounds (song, bang, beep, rattle, snore), views, smells, appetizing, motivation. Opposed to other causes such as Agents, Forces, Chance.

Phenomenal

Situations that occur in nature controlled or uncontrolled or considered as a force; e.g. weather, chance. Opposed to other causes such as Stimuli, Agents.

As far as the set of Base Concepts is representative for the total wordnets, this set of SituationComponents is also representative for the whole. Note that adjectives and adverbs have not been classified in EuroWordNet yet. In this respect we may need a further elaboration of these components when these parts-of-speech are added. The last three SituationComponents are subdivided, which are discussed in the following subsections.

As said above, a verb or 2ndOrder noun may thus be composed of any combination of these components. However, it is obvious that some combinations make more sense than others. Situations involving Purpose often also involve Cause, simply because it is in the nature of our behavior that people do things for some purpose. Furthermore, there may be some specific constraints that some components are restricted to some SituationTypes. Cause and Purpose can only occur with Dynamic Situations. When there is no constraint we will thus get various combinations, such as Dynamic and Physical for *to colour* or Static and Physical for *colour*, where word meanings can still be grouped on the basis of the shared component: Physical.

The more specific a word is the more components it incorporates. Just as with the 1stOrderEntities we therefore typically see that the more frequent classifying nouns and verbs only incorporate a few of these components. In the set of common Base-Concept, such classifying words are more frequent, and words with many SituationComponents are therefore rare. In Appendix II a list is given of al TC combinations with the clusters of BCs that belong to it. Appendix III gives a list of all cluster combinations with frequency. The 1stOrderEntities (491 BCs) are divided over 124 clusters, the 2ndOrderEntities (500 BCs) over 314 clusters.

Finally, it is important to realize that the Top Ontology does not necessarily correspond with the language-internal hierarchies. Each language-internal structure has a different mapping with the top-ontology via the ILI-records to which they are linked as equivalences. For example there are no words in Dutch that correspond with technical notions such as 1stOrderEntity, 2ndOrderEntity, 3rdOrderEntity, but also not with more down-to-earth concepts such as the Functional 1stOrder concept Container. These levels will thus not be present in the Dutch wordnet. From the Dutch hierarchy it will hence not be possible to simply extract all the *containers* because no Dutch word meaning is used to group or classify them. Nevertheless, the Dutch ‘containers’ may still be found either via the equivalence relations with English ‘containers’ which are stored below the sense of “container” or via the TopConcept clustering Container that is imposed on the Dutch hierarchy (or any other ontology that may be linked to the ILI). See Peters et al. (1998) for a further discussion on accessing the different modules in the database.

The Top-Concepts have been assigned directly to the Base Concepts but also to other tops in WordNet1.5 that are not included in the Base Concept selection (389 verbal synsets and 2 nominal synsets). This resulted in 793 nominal and 617 verbal synsets that have been classified in total. The file with these classifications is provided on the general EuroWordNet CD and can be downloaded from the WWW-site.

By inheriting these top-concept assignments via the hyponymy relations it is possible to populate the complete ILI with top-concepts. However, because we want to keep a distinction between the directly assigned and the inherited top-concepts we decided to add the inherited top-concepts to the glosses. There are two things to be noted with respect to the inherited top-concepts. First of all, redundant top-concepts are added in so far they have not been inherited from higher levels. If a top-concept list includes Animal but not Natural, then Natural is added because it is implied by Animal according to the above top-concept hierarchy. The second point is that the hyperonym classification of WordNet1.5 is not always the same or consistent with our top-ontology assignement. This can be a matter of choice, because we did not agree with theWordNet1.5 classification or it may be incidental because top-concepts, assigned to the higher levels, are no longer valid at deeper levels of the hierarchy. Examples of the former case are 3rdOrderEntities that have been classified in WordNet1.5 below *psychological_feature* that goes to *state* together will all statitive nominals. In EuroWordNet, states are static 2ndOrderEntities and the WordNet1.5 top *state* has been classified accordingly. Consequently, many 3rdOrderEntities will thus inherit both the top-concepts 2ndOrderEntity and 3rdOrderEntity. Inconsistencies at lower levels, the second possibility of mismatch, may arise. We have not been able to verify the inherited top-concepts at all levels.

Finally, we have added the lexicographer's file codes in WordNet1.5 to the glosses as well. Since these are assigned on a synset to synset basis, it was not necessary to inherit these codes. The compatibility of the lexicographer's file-codes and the top-ontology is given below in 16. Below are some examples of ILI-record glosses that include the augmented the lexigrapher's file code and the inherited EuroWordNet top-concepts (where redundant TCs are added as well):

```
0 ILI_RECORD
  1 PART_OF_SPEECH "n"
  1 FILE_OFFSET 2728
  1 GLOSS "any living entity& 03 1stOrderEntity Living Natural Origin"
  1 VARIANTS
    2 LITERAL "life form"
      3 SENSE 1
    2 LITERAL "organism"
      3 SENSE 1
    2 LITERAL "being"
      3 SENSE 1
    2 LITERAL "living thing"
      3 SENSE 1

0 ILI_RECORD
  1 PART_OF_SPEECH "n"
  1 FILE_OFFSET 1978911
  1 GLOSS "a flat-bottomed boat used on upper Great Lakes& 03 06 1stOrderEntity
Artifact Form Function Instrument Object Origin Vehicle"
  1 VARIANTS
    2 LITERAL "Mackinaw boat"
      3 SENSE 1
    2 LITERAL "mackinaw"
      3 SENSE 2
```

```
0 ILI_RECORD
  1 PART_OF_SPEECH "v"
  1 FILE_OFFSET 1064210
  1 GLOSS "roll around, as of a pig in mud& 2ndOrderEntity 38 Dynamic Location
Physical SituationType"
  1 VARIANTS
    2 LITERAL " roll around "
      3 SENSE 1
    2 LITERAL " wallow "
      3 SENSE 2
    2 LITERAL " welter"
      3 SENSE 3
```

Table 17: Mapping of WordNet1.5 Lexicographer's file codes to EuroWordNet top-concepts

Code	*WordNet File Name*	*EuroWordNet Top Concepts*
03	noun.Tops	
04	noun.act	Agentive;
05	noun.animal	Animal;
06	noun.artifact	Artifact;
07	noun.attribute	Property;
08	noun.body	Object; Natural;
09	noun.cognition	Mental;
10	noun.communication	Communication;
11	noun.event	Dynamic;
12	noun.feeling	Experience;
13	noun.food	Comestible;
14	noun.group	Group;
15	noun.location	Place;
16	noun.motive	3rdOrderEntity;
17	noun.object	Object;
18	noun.person	Human;
19	noun.phenomenon	Phenomenal;
20	noun.plant	Plant;
21	noun.possession	Possession;
22	noun.process	Dynamic;
23	noun.quantity	Quantity;
24	noun.relation	Relation;
25	noun.shape	Physical;
26	noun.state	Static;
27	noun.substance	Substance;
28	noun.time	Time;
29	verb.body	Dynamic; Physical;
30	verb.change	Dynamic;
31	verb.cognition	Mental; Dynamic;
32	verb.communication	Communication; Dynamic;
33	verb.competition	Social; Dynamic;
34	verb.consumption	Physical; Location; Dynamic;
35	verb.contact	Location; Dynamic;
36	verb.creation	Existence; BoundedEvent;
37	verb.emotion	Experience; Mental;
38	verb.motion	Location; Physical; Dynamic;
39	verb.perception	Experience; Physical; Dynamic;
40	verb.possession	Possession; Dynamic;
41	verb.social	Social; Dynamic;
42	verb.stative	Static;
43	verb.weather	Phenomenal; Physical; Dynamic;

The EuroWordNet database

The multilingual EuroWordNet database consists of three components:

1. The actual wordnets in Flaim database format: an indexing and compression format of Novell, which is also part of the Groupwise software.
2. Polaris (Louw 1998): a wordnet editing tool for creating, editing and exporting wordnets.
3. Periscope (Cuypers and Adriaens 1997): a graphical database viewer for viewing and exporting wordnets.

The Polaris tool is a re-implementation of the Novell ConceptNet toolkit (Díez-Orzas et al 1995) adapted to the EuroWordNet architecture. Polaris can import new wordnets or wordnet fragments from ASCII files with the correct import format and it creates an indexed EuroWordNet database (an example of the import format is the Top Ontology file). Furthermore, it allows a user to edit and add relations in the wordnets and to formulate queries. The Polaris toolkit makes it possible to visualize the semantic relations as a tree-structure that can directly be edited. These trees can be expanded and shrunk by clicking on word-meanings and by specifying so-called TABs indicating the kind and depth of relations that need to be shown, see Figure 13 below. Expanded trees or sub-trees can be stored as a set of synsets, which can be manipulated, saved or loaded. Additionally, it is possible to access the ILI or the ontologies, and to switch between the wordnets and ontologies via the ILI. Finally, it contains a query interface to match sets of synsets across wordnets. This can be down in several general ways:

1. multiple windows that expand separate wordnets and show the equivalence relations (see Figure 13)
2. looking up inter-lingual-index items (Explore ILI-records) which will give the associated synsets in each language (see Figure 14)
3. looking up Top-Concepts, which will give associated ILI-records (mostly Base Concepts) and the synsets in each language that are associated with these (see Figure 15)
4. looking up Domains, which will give associated ILI-records (mostly more specific concepts) and the synsets in each language that are associated with these (see Figure 16)
5. projecting a set of synsets in one language to a target language, via a selected set of equivalence relations (see Figure 17).

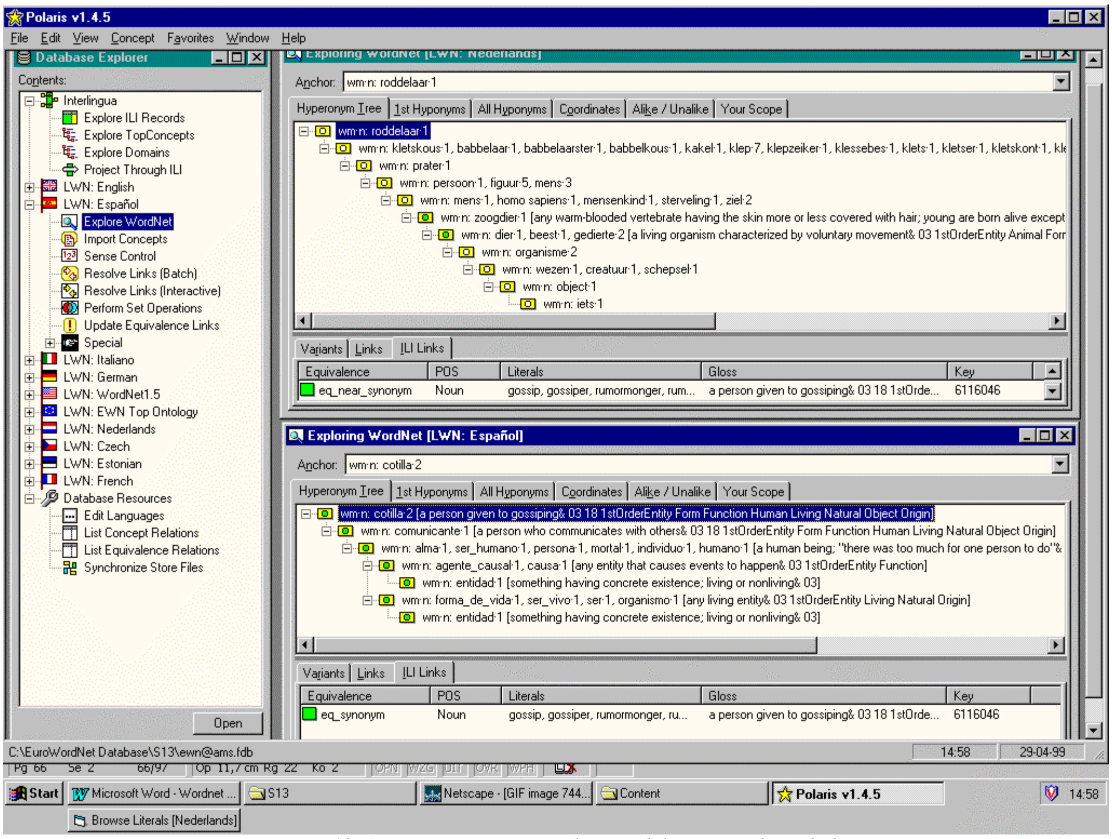


Figure 13: Accessing separate wordnets and their equivalence links

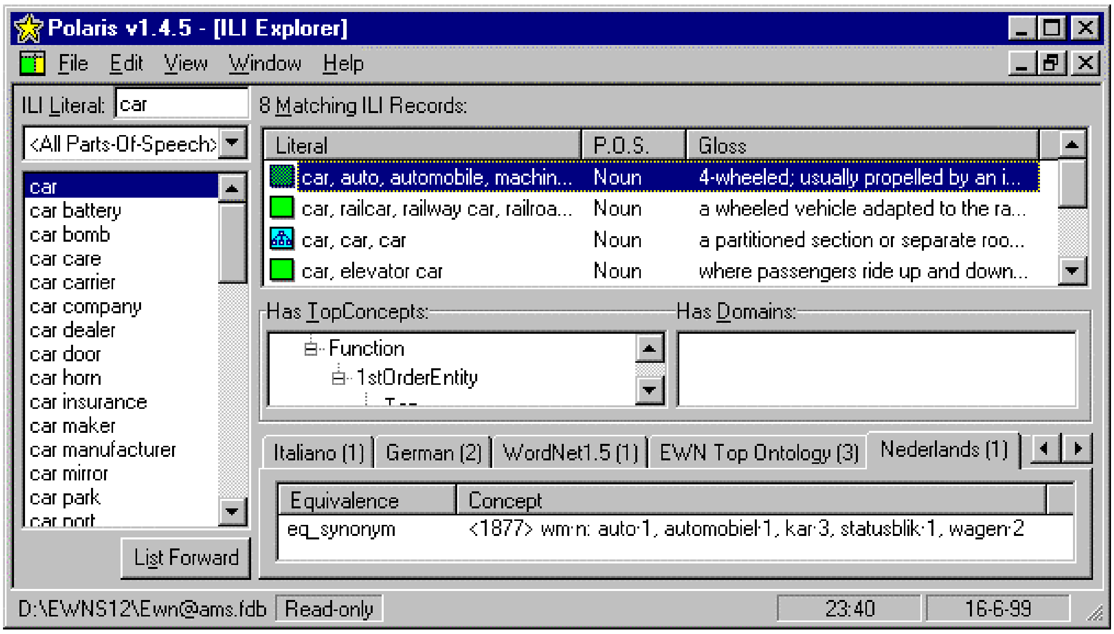


Figure 14: Accessing different wordnets via the Inter-Lingual-Index

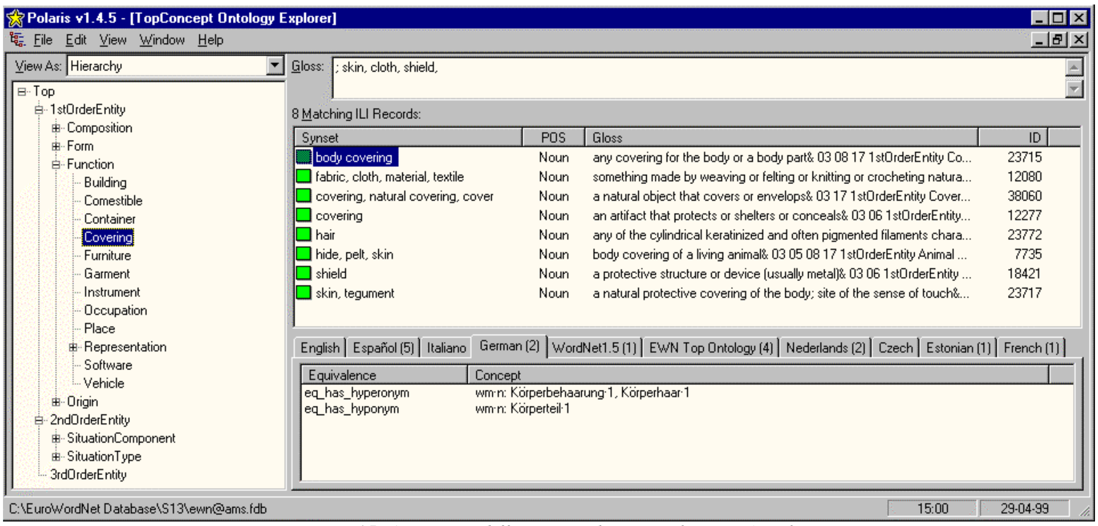


Figure 15: Accessing different wordnets via the Top-Ontology

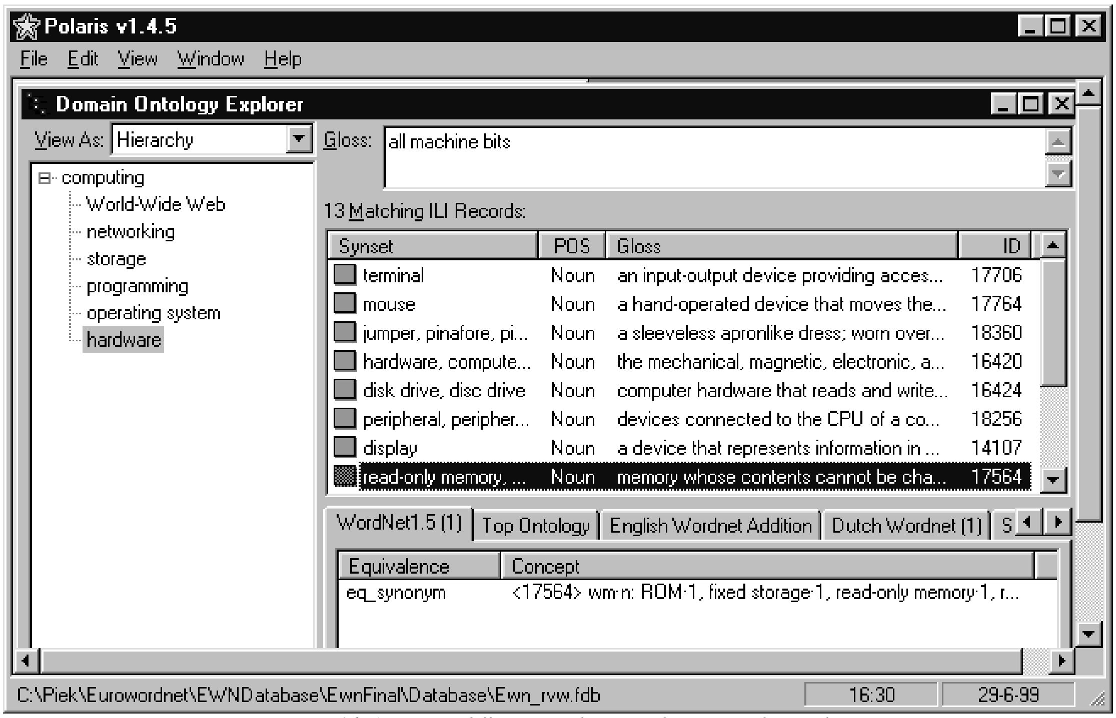


Figure 16: Accessing different wordnets via the Domain hierarchy

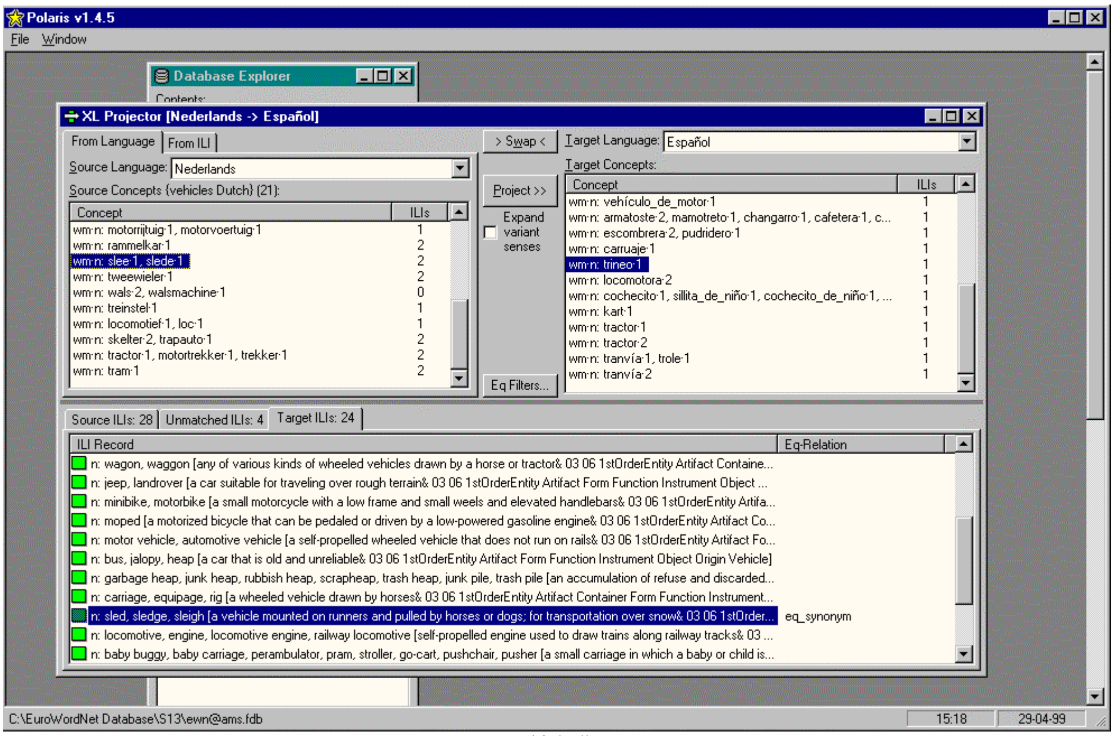


Figure 17: Projecting Dutch "vehicles" *(1 level) to the Spanish wordnet*

In the case of a projection, which is shown in Figure 16, a selection of synsets in a particular language (as shown in the left upper window for Dutch vehicles) is loaded and the desired types of equivalence mapping are selected. When a target language is choosen, the ILI-records that match the equivalence types are taken to generate the synsets in the target language also linked to them. The resulting set of target synsets is given in the right upper window, as is shown here for Spanish. The lower window gives, with different TABs, the ILI-records that are linked in the source selection, the ILI-records that could not be matched and the records that are shared by the source and target.

The cross links can also be activated by double-clicking the synsets or the ILI-records. For example, double-clicking a ILI-record that is given as an equivalent for a synset in the language-specific explorer, will activate the ILI-explorer and from there it is possible to select a synset in another language.

The Periscope program is a public viewer that can be used to look at wordnets created by the Polaris tool and compare them in a graphical interface. Word meanings can be looked up and trees can be expanded. Individual meanings or complete branches can be projected on another wordnet or wordnet structures can be compared via the equivalence relations with the Inter-Lingual-Index. Selected trees can be exported to Ascii files. The Periscope program cannot be used for importing or changing wordnets. Examples of the Periscope interface have already been given in this document.

5. Description of the CD-Rom

The EuroWordNet results are distributed by ELRA/ELDA. The distribution consists of:

1. a general CD containing all the freeware and public data (also includes this document)
2. for each language: a language specific CD

The Polaris wordnet toolkit should be licensed from Lernout and Hauspie. Contact person is Geert Adriaens (e-mail: Geert.Adriaens@lhs.be).

The General CD contains the following data:
DOC:
- EuroWordNet General Documentation (this document)
    EWN GENERAL.ps (PostScript),
    EWN GENERAL.doc (Word-97).
    EWN GENERAL.html
- EuroWordNet Powerpoint Presentation
    EWN GENERAL.ppt

- Text data
    - BaseConcepts:
        The Base Concepts with top-concept clasification and WordNet1.5 classification
        - NOUN_BASECONCEPTS.txt & VERB_BASECONCEPTS.txt
    - Inter-Lingual-Index:
        - ILI_WN15.ewn (ILI based on WordNet1.5)
        - ILI_CLUSTERS.ewn (added composite ILI-records or clusters)
        - ILI_DOMAIN_LABELS.ewn (Domain labels assigned to ILI)
        - ILI_TOP_ONTOLOGY.ewn (Top Concepts assigned to ILI)
        - ILI_COMPUTER_TERMS.ewn (computer terminology added and glossed)
    - WordNet15:
        WordNet1.5 in EWN format:
        - WN_15_nouns.ewn, WN_15_verbs.ewn,
        - WN_15_adjectives.ewn, WN_15_adverbs.ewn
    - Samples:
        EuroWordNet Samples in EWN format:
        - WN_NL.ewn, WN_IT.ewn, WN_ES.ewn, WN_DE.ewn,
        - WN_FR.ewn, WN_EE.ewn, WN_CZ.ewn, WN_TO.ewn (top-ontology as wordnet)
- EwnDataBase:
    The EuroWordNet database with the ILI and separate stores (*.sdb) for
    the wordnet samples and WordNet1.5 (see Figure 18 below).
- PERISCOPE:
    - Periscope software, to be installed on Windows95/98/NT
    - MAN: Periscope manual and installation instructions

- Readme.txt

Explanations:
*.sdb = Polaris database format;
*.ewn = EuroWordNet format that is exported by Polaris and can be imported by it;

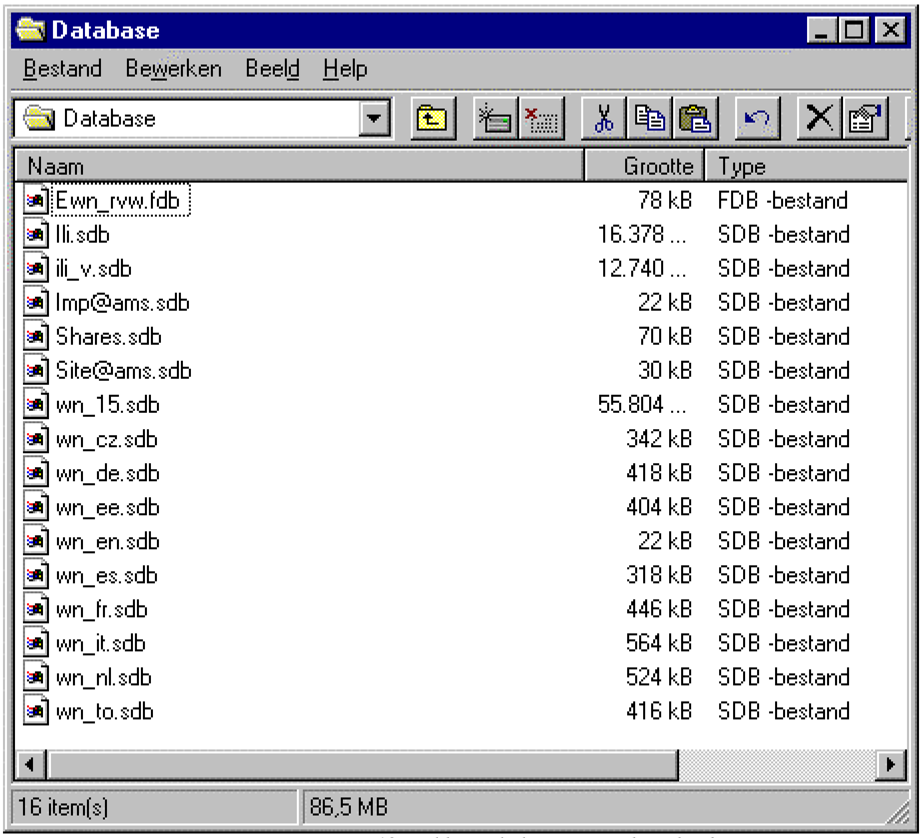


Figure 18: Folder with the EuroWordNet database stores

Figure 18 lists the database files that should be present for Periscope and Polaris to operate properly. Only the file “Ewn_rvw.fdb” should be opened by both programs after launch. The individual wordnets are stored as the *.sdb files. Note that the top-ontology is also included as a mini-wordnet so that it can be accessed in Periscope. Polaris also has an integrated version of the top-ontology.

For each language there will be a language-specific CD which contains:

- language.sdb (the complete database accessible by Periscope or Polaris)
- language.ewn (ascii version of the database in EWN import/export format)
- document on the content of the wordnet (Postscript, Word-97)
- document on the comparison of the wordnets (Postscript, Word-97)

The content documentation includes a description of the individual wordnets and a comparison of them. This comparison document is released separately for EuroWordNet-1 (LE2-4003) and EuroWordNet-2 (LE4-8328). The former includes descriptions of the English, Dutch, Spanish and Italian wordnets and a comparison of these. The latter includes a description of the French, German, Estonian and Czech and their comparison. These documents can also be downloaded from the EuroWordNet WWW-site.

The general CD is distributed in addition to one or more language-specific CDs. A user can then replace the language-sample.sdb (keep a copy in a separate folder!) by the full language.sdb file and directly see it with Periscope. In this way, it is not necessary to make a *.fdb for this language (or any combination of languages) with Polaris, and it thus is not necessary to buy Polaris before one can see the database. If languages are missing in the folder Periscope does not work (and also Polaris may crash). So make sure that a copy of each of the language.sdb files is present in the database folder.

Acknowledgements

The EuroWordNet project has been possible by the funding of the European Commision, DGXIII, Luxembourg. We also want to thank the publishers Van Dale Lexicografie and Bibliograf who have provided their material. We clearly want to express our gratitude to the Princeton research group of George Miller, for building wordnet, for inspiring us and for giving us the possibility to build the wordnets in the other languages. Without the Princeton wordnet, there would not have been EuroWordNet.

For the rest, we want to thank all people that contributed to the project.

References

Ageno A., F. Ribas, G. Rigau, H. Rodriquez and F. Verdejo. 1993. *TGE: Tlinks Generation Environment*. Acquilex II (BRA 7315) Working Paper 7. Polytecnica de Catalunya, Barcelona.

Agirre E. and Rigau G. 1996. Word Sense Disambiguation using Conceptual Density, in proceedings of the 16th International Conference on Computational Linguistics (COLING'96). Copenhagen, Denmark.

Alonge, A. 1996. *Definition of the Links and Subsets for Verbs*. Deliverable D006, *EuroWordNet*, LE2-4003, Computer Centrum Letteren, University of Amsterdam.

Alonge, A., N. Calzolari, P. Vossen, L. Bloksma, I. Castellon, T. Marti, W. Peters. 1998. "The Linguistic Design of the EuroWordNet Database". In: Nancy Ide, Daniel Greenstein, Piek Vossen (eds), *Special Issue on EuroWordNet. Computers and the Humanities*, Volume 32, Nos. 2-3 1998. 91-115.

Alvar M. (Ed.) 1987 *Diccionario General Ilustrado de la Lengua Española VOX*. Biblograf S.A. Barcelona.

Apresjan, J. 1973. "Regular Polysemy". *Linguistics* 142.

Bateman, J., B. Magnini and J. Rinaldi 1994 *The Generalised Upper Model*. In Proceedings of ECAI 1994.

Beckwith, R. and G.A. Miller. 1990. “Implementing a Lexical Network”. *International Journal of Lexicography,* Vol 3, No.4, 302-312.

Buitelaar, P. 1998. *Corelex: Systematic Polysemy and Underspecification*, PhD., Department of Computer Science, Brandeis University.

Climent, S., H. Rodriguez, J. Gonzalo. 1996. *Definition of the Links and Subsets for Nouns of the EuroWordNet Project*. Deliverable D005, EuroWordNet, LE2-4003, Computer Centrum Letteren, University of Amsterdam.

Copeland, C., Durand, J., Krauwer, S. and Maegaard, B. (eds.). 1991. *The Eurotra Formal Specifications*, Office for Official Pblications of the European Community, Luxembourg.

Copestake A. and A. Sanfilippo. 1993. *Multilingual Lexical Representation*, Acquilex II (BRA 7315) Working Paper 2. Cambridge University.

Copestake A. and Briscoe, T. 1991. "Lexical operations in a unification-based framework". Ed. Pustejovsky J. and Bergler S. *Lexical Semantics and Knowledge Representation,* Association for Computational Linguistics.

Copestake A., T Briscoe, P. Vossen, A Ageno, I Castellon, F Ribas, G Rigau, H Rodriguez, A Sanmiotou. 1995. "Acquisition of Lexical Translation Relations from MRDs". *Journal of Machine Translation*, Volume 9, issue 3.

Copestake, A. 1995. "Representing Lexical Polysemy". *Proceedings of AAAI,* Stanford Spring Symposium, Stanford.

Cruse, D. A. 1986. *Lexical Semantics*. Cambridge, Cambridge University Press.

Cuypers, I. And G. Adriaens. 1997. *Periscope: the EWN Viewer*, EuroWordNet Project LE4003, Deliverable D008d012. University of Amsterdam, Amsterdam.

Díez Orzas, P., M. Louw and Ph. Forrest. 1996. *High level design of the EuroWordNet Database*. EuroWordNet Project LE2-4003, Deliverable D007.

Díez-Orzas P. and I. Cuypers. 1995. *The Novell ConceptNet*, Internal Report, Novell Belgium NV.

Dik, S. *Stepwise Lexical Decomposition*. Lisse, Peter de Ridder Press, 1978.

Donker, T. , I. Serail and P. Vossen, 1994, “Salient Words and Phrases”, Computer Centrum Letteren, University of Amsterdam, March 1994, Sift deliverable D5, LRE 62030.

Dowty, D.R. 1979. Word meaning and Montague grammar: the Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ. Dordrecht: Reidel.

Dowty, D., On the Semantic Content of the Notion of Thematic Role, in G. Cherchia, B. Partee, R. Turner (eds), Properties, Types and meaning, Kluwer, 1989.

Fellbaum, C. (ed.) 1998. *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Fellbaum, C. 1990. “English Verbs as a Semantic Net”. *International Journal of Lexicography,* Vol 3, No.4, 278-301.

Gross, D. and K.J. Miller 1990. “Adjectives in Wordnet”. *International Journal of Lexicography*, Vol 3, No.4, 265-277.

Gruber, T.R. 1992 *Ontolingua: a Mechanism to Support Portable Ontologies*. Report KSL 91-66. Stanford University.

Lakoff, G. 1987. Women, Fire and Dangerous Things, University of Chicago Press, Chicago/London.

Lenat, D. and R. Guha 1990 *Building Large Knowledge-based Systems. Representation and Inference in the CYC Project.* Addison Wesley.

Levin, B. 1993 *English Verb Classes and Alternations*. University of Chicago Press. Chicago.

Louw, M. 1998. *The Polaris User's Guide: The EuroWordNet Database Editor.* EuroWordNet (LE4-4003 Deliverable D024*)* University of Amsterdam..

Lyons, J. 1977. *Semantics*. London, Cambridge University Press, 1977.

McCarthy, D. 1997. "Word sense disambiguation for acquisition of selectional preferences." Ed. Vossen, P., Adriaens, G., Calzolari, N., Sanfilippo, A., and Wilks, Y., *Proceedings of the ACL/EACL'97 Workshop on Automatic Information Extraction and Building of Lexical Semantic Resources*.

Miller G., R. Beckwith, C. Fellbaum, D. Gross, K.J. Miller. 1990. “Introduction to WordNet: An On-line Lexical Database". *International Journal of Lexicography*, Vol 3, No.4, 235-244.

Nirenburg, S. (ed.). 1989. "Knowledge-based MT", *Special issue Machine Translation* vol.4, no 1 and 2, Kluwer Publishers, Dordrecht.

Nunberg, G & A. Zaenen. 1992. "Systematic Polysemy in Lexicology and Lexicography". *Proceedings of EURALEX’92*, University of Tampere.

Ostler, N. and S. Atkins. 1991. "Predictable Meaning Shift: some linguistic properties of lexical implication rules". Ed. Pustejovsky J. and Bergler S. *Lexical Semantics and Knowledge Representation*, Association for Computational Linguistics.

Peters, W., I. Peters, and P. Vossen. 1998a. "The Reduction of Semantic Ambiguity in Linguistic Resources". In: A. Rubio, N. Gallardo, R. Catro and A. Tejada (ed) *Proceedings of First International Conference on Language Resources and Evaluation*, Granada, 28-30 May 1998. 409-416

Peters, W., P. Vossen, P. Diez-Orzas, G. Adriaens. 1988b. "Cross-linguistic Alignment of Wordnets with an Inter-Lingual-Index. In: Nancy Ide, Daniel Greenstein, Piek Vossen (eds), *Special Issue on EuroWordNet. Computers and the Humanities*, Volume 32, Nos. 2-3 1998. 221-251.

Peters, W. 1999. The Inter-Lingual-Index in EuroWordNet, EuroWordNet Deliverable 2D004, Project reference LE4-8328, University of Amsterdam.

Procter, P (ed) 1978 *Longman Dictionary of Contemporary English.* Longman, Harlow and London.

Pustejovsky, J. 1991. The syntax of event structure, Cognition, 41, 47-81.

Pustejovsky, J. 1995. *The Generative Lexicon*, MIT Press, Cambridge MA.

Rodriquez, H., S. Climent, P. Vossen, L. Bloksma; A. Roventini, F. Bertagna, A. Alonge, W. Peters. 1998. "The Top-Down Strategy for Building EuroWordNet: Vocabulary Coverage, Base Concepts and Top Ontology". In: Nancy Ide, Daniel Greenstein, Piek Vossen (eds), *Special Issue on EuroWordNet. Computers and the Humanities*, Volume 32, Nos. 2-3 1998. 117-152.

Rosch, E. 1977 *Human Categorisation.* In N. Warren (Ed.) Studies in Cross-Cultural Psychology, Vol. I, pp. 1-49. Academic Press. London.

Rosch, E. 1975. Cognitive Representations of Semantic Categories In: Journal of Experimental Psychology: General 104:192-233.

Sanfilippo, A. 1997. "Using semantic similarity to acquire co-occurrence restrictions from corpora". Ed. Vossen, P., Adriaens, G., Calzolari, N., Sanfilippo, A., and Wilks, Y., *Proceedings of the ACL/EACL'97 Workshop on Automatic Information Extraction and Building of Lexical Semantic Resources*.

Sanfilippo, A., T. Briscoe, A. Copestake, M. A. Martì-Antonin, A. Alonge. 1992. “Translation Equivalence and Lexicalization in the ACQUILEX LKB.” *Proceedings of the 4th International Conference on Theoretical and Methodological Issues in Machine Translation*, Montreal, Canada.

Talmy, L. 1985. “Lexicalization Patterns: Semantic Structure in Lexical Form.” In *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon*. Ed. T. Shopen. Cambridge, Cambridge University Press.

Vendler, Z. 1967. Linguistics and philosophy. Ithaca: Cornell University Press.

Verkuyl, H. 1972. On the compositional nature of the aspects. Dordrecht: Reidel.

Verkuyl , H. 1989. “Aspectual classes and aspectual distinctions”, Linguistics and Philosiphy, 12, 39-94.

Vossen P. and A. Bon 1996 Building a semantic hierarchy for the Sift project, Sift LRE 62030, Deliverable D20b, University of Amsterdam. Amsterdam.

Vossen, P., L. Bloksma, C. Peters, A. Alonge, A. Roventini, E. Marinai, I. Castellon, T. Marti, G. Rigau, 1998. "Compatibility in Interpretation of Relations in EuroWordNet". In: Nancy Ide, Daniel Greenstein, Piek Vossen (eds), *Special Issue on EuroWordNet. Computers and the Humanities*, Volume 32, Nos. 2-3 1998. 153-184.

Vossen P. (eds.) 1999 EuroWordNet: a multilingual database with lexical semantic networks for European Languages. Kluwer Academic Publishers, Dordrecht.

Vossen P. 1999 fc., EuroWordNet as a multilingual database. In: Wolfgang Teubert (ed), Mouton Gruyter, Berlin.

Vossen, P. , W. Peters, J. Gonzalo, 1999, Towards a Universal Index of Meaning, In proceedings of the Siglex99-workshop, ACL-1999, Maryland.

VOX-HARRAP’S 1992. *Diccionario Esencial Español-Inglés Inglés-Español*. Biblograf S.A. Barcelona.

Winston M., R. Chaffin, D. Herrmann. 1987. “A Taxonomy of Part-Whole Relations.” *Cognitive Science* 11, 417–444.

Appendix I: Base Concepts Selected by four sites in EuroWordNet

NOMINAL BASE CONCEPTS SELECTED BY ALL FOUR SITES

act 1*	element 6	ornament 1
activity 1	fabric 1	period 3
amount of time 1	fauna 1	period of time 1
animal 1	feeling 1	person 1
animate being 1	flora 1	phenomenon 1
attitude 3	food 1	plant 1
beast 1	ground 7	plant life 1
beverage 1	human 1	point 12
brute 1	human action 1	potable 1
chemical compound 1	human activity 1	quality 1
chemical element 1	individual 1	solid ground 1
cloth 1	knowledge 1	someone 1
cognition 1	land 6	soul 1
compound 4	line 26	structure 1
construction 4	material 1	stuff 7
creature 1	material 5	substance 1
decoration 2	matter 1	terra firma 1
drink 2	mental attitude 1	textile 1
dry land 1	mortal 1	time period 1
earth 3	nutrient 1	worker 2

Verbal Base Concepts selected by all four sites

be 4	have 7	move 15
cause 6	have the quality of being 1	remove 2
cover 16	induce 2	stimulate 3
create 2	locomote 1	take 4
get 9	make 12	take away 1
go 14	make 13	travel 4

*Sense numbers do not necessarily correspond with the sense numbers in WordNet1.5

83

Appendix II Top Ontology Classification of the Base Conceps

1stOrderEntity

thing 2: 01958400-n
Artifact
article 1: 00012356-n
Building+Group+Artifact
establishment 2: 01960381-n
Building+Group+Object+Artifact
factory 1: 02895948-n
housing 3: 02724446-n
Building+Object
abode 1: 02456156-n
Building+Object+Artifact
building 3: 02207842-n
building complex 1: 02209583-n
business establishment 1: 01960698-n
house 2: 02728393-n
mercantile establishment 1: 01961354-n
plant 2: 02893856-n
shop 1: 03066446-n
Building+Part+Object+Artifact
office 4: 01960921-n
room 1: 02725092-n
Comestible
aliment 1: 04837708-n
condiment 1: 05019688-n
dainty 1: 04856504-n
Comestible+Artifact
baked good 1: 04875085-n
candy 1: 04859051-n
course 5: 04842977-n
dish 3: 04843172-n
Comestible+Group+Artifact
pastry 2: 04875625-n
Comestible+Group+Plant
garden truck 1: 04935405-n
Comestible+Liquid
beverage 1: 05074818-n
drink 4: 05077192-n
Comestible+Liquid+Artifact
alcohol 2: 05076795-n
sauce 1: 05034282-n
vino 1: 05081539-n
Comestible+Object+Plant
edible fruit 1: 04935607-n
vegetable 1: 04937211-n
Comestible+Part
helping 2: 04842062-n
ingredient 3: 05018259-n
Comestible+Part+Solid
commissariat 1: 04838667-n
Comestible+Part+Solid+Natural
herb 1: 05020240-n
Comestible+Solid+Animal
meat 2: 04894971-n
Comestible+Solid+Artifact
bread 1: 04916628-n
cake 2: 04879808-n
cheese 1: 05050320-n
dessert 1: 04867005-n
refined sugar 1: 05056815-n
Comestible+Substance
comestible 1: 04830190-n
dairy product 1: 05045392-n
flavorer 1: 05018491-n
food 1: 00011263-n
foodstuff 2: 04834499-n
Comestible+Substance+Artifact
confection 2: 04858776-n
Container+Object
container 1: 01990006-n
vessel 2: 03236256-n
Container+Object+Artifact
bottle 1: 02180350-n
tube 2: 03219464-n
Container+Part+Solid+Living
blood vessel 1: 03733773-n
passage 7: 03622270-n
tube 4: 03621461-n
vas 1: 03725681-n
vein 2: 03734105-n
Container+Solid
channel 1: 02342911-n
passage 6: 02857000-n
Container+Solid+Artifact
bag 4: 02097669-n
Covering
shield 2: 02895122-n
Covering+Artifact
covering 4: 01991765-n
Covering+Object+Natural
cover 7: 05639760-n
Covering+Part+Solid+Living
body covering 1: 03616903-n
hair 2: 03626404-n
skin 4: 03617358-n
Covering+Part+Solid+Natural
hide 1: 01246669-n
Covering+Solid+Artifact
cloth 1: 01965302-n
Creature
deity 1: 05774165-n
imaginary being 1: 05764486-n
Function
Function
asset 2: 08179398-n
barrier 1: 02117075-n
belonging 2: 08128156-n
building material 1: 08885624-n
causal agency 1: 00004473-n
commodity 1: 02329807-n

consumer goods 1: 02344541-n
creation 3: 01992919-n
curative 1: 02024781-n
decoration 2: 02029323-n
device 4: 04576638-n
fastener 1: 02494190-n
force 6: 06276483-n
force 7: 06491991-n
form 5: 03957219-n
impediment 1: 02822812-n
medicament 1: 02011101-n
possession 1: 00017394-n
protection 4: 02937777-n
remains 2: 05638634-n
restraint 2: 02995085-n
support 6: 03149538-n
support 7: 03150440-n
supporting structure 1: 03150653-n

Function+Artifact
art 2: 02980374-n
facility 1: 01962758-n
piece of work 1: 02932267-n
plaything 1: 02032220-n
product 2: 02929839-n
thing 3: 01958716-n

Function+Group+Human
church 3: 05168576-n
club 6: 05238189-n
company 2: 05218109-n
company 3: 05220757-n
educational institution 1: 05270729-n
establishment 4: 05152219-n
house 6: 05206050-n
house 8: 05236426-n
institute 1: 05334108-n
organization 5: 05149489-n
party 3: 05259394-n
school 5: 05271053-n
state 3: 05214009-n
union 7: 05286371-n

Function+Living
reproductive structure 1: 06668106-n

Function+Object+Artifact
card 1: 02245777-n
painting 4: 02985557-n

Function+Object+Human
defender 1: 05844515-n
negotiant 1: 06224003-n
representative 3: 06305438-n

Function+Part+Object+Artifact
grip 3: 02598444-n

Function+Solid+Natural
ground 6: 05719829-n

Function+Substance
combustible 1: 08936946-n
cushioning 1: 02841356-n
Functional
means 2: 02766526-n

Furniture+Group+Artifact
furnishings 2: 02043015-n

Furniture+Object+Artifact
article of furniture 1: 02008299-n
chair 2: 02275608-n
seat 2: 03044397-n
table 1: 03160216-n
table 2: 03160884-n

Garment+Solid+Artifact
apparel 1: 02307680-n
garment 1: 02309624-n
headdress 1: 02612319-n
Gas
gas 5: 08938440-n
Group
accumulation 2: 05120211-n
arrangement 7: 05114274-n
group 1: 00017008-n
set 7: 05142366-n
system 1: 02036726-n
system 7: 05354739-n
unit 1: 01959683-n

Group+Human
a people 1: 05208026-n
administration 3: 05207180-n
administrative unit 1: 05233375-n
agency 1: 05301461-n
assemblage 4: 05132844-n
association 3: 05150995-n
authorities 1: 05151482-n
band 7: 05246785-n
body 7: 05127029-n
body politic 1: 05209013-n
citizenry 1: 05205244-n
commission 7: 05293372-n
community 2: 05236204-n
company 1: 05217925-n
division 9: 05233198-n
enterprise 3: 05154048-n
family 2: 05129983-n
family 3: 05131472-n
hoi polloi 1: 05214761-n
human race 1: 05116306-n
movement 7: 05365815-n
party 2: 05255204-n
people 1: 05116476-n
populace 1: 05214471-n
social group 1: 05119847-n
unit 4: 05222733-n

Group+Living
life 1: 00003504-n

Group+Plant
flora 1: 00008894-n
ImageRepresentation
figure 12: 08483587-n
line 26: 08484352-n

ImageRepresentation+Artifact
design 2: 02030692-n
emblem 2: 04481847-n
icon 1: 02879254-n
representation 3: 02354709-n

ImageRepresentation+Object

solid 1: 08482581-n
ImageRepresentation+Object+Artifact
art 4: 04539476-n
bill 7: 04427449-n
Instrument+Artifact
equipment 1: 02004554-n
instrumentality 1: 02009476-n
light 1: 02697378-n
mechanism 2: 02010561-n
Instrument+Group
material 2: 02765238-n
Instrument+Group+Object+Artifact
arm 4: 03253503-n
arms 2: 03254035-n
Instrument+Object+Artifact
apparatus 1: 02069513-n
device 2: 02001731-n
engine 1: 02473560-n
implement 1: 02008805-n
instrument 2: 02657448-n
machine 2: 02743730-n
machine 3: 02744991-n
measuring instrument 1: 02766721-n
motor 1: 02798554-n
musical instrument 1: 02804379-n
tool 2: 03198235-n
LanguageRepresentation
alphabetic character 1: 04451043-n
appellation 1: 04183149-n
language 3: 04155501-n
language unit 1: 04156286-n
message 1: 04139704-n
natural language 1: 04495739-n
word 1: 04157535-n
LanguageRepresentation+Artifact
character 5: 04444555-n
document 2: 04242515-n
document 3: 08225885-n
identification number 1: 04230965-n
letter 1: 04330686-n
literary composition 1: 04196450-n
mark 8: 04443464-n
material 3: 04197046-n
name 1: 04180885-n
number 7: 04435360-n
poem 1: 04203578-n
printed symbol 1: 04443305-n
publication 3: 04308479-n
register 5: 08232464-n
text 1: 04211005-n
title 2: 04183413-n
writing 4: 04195435-n
written communication 1: 04187642-n
LanguageRepresentation+Group+Artifact
line 15: 04547144-n
LanguageRepresentation+Object+Artifact
book 3: 02675934-n
book 5: 04222100-n
book of facts 1: 04226531-n
record 6: 08226179-n
LanguageRepresentation+Part+Artifact
end 4: 03973920-n
LanguageRepresentation+Part+Object+Artifac
t
issue 5: 04312465-n
LanguageRepresentation+Solid+Artifact
bill of fare 1: 04253617-n
symbolic representation 1: 04192746-
n
Liquid
acid 2: 08796177-n
fluid 1: 08975815-n
fluid 2: 08976164-n
lipid 1: 08975312-n
liquid 4: 08976498-n
oil 2: 08991530-n
Living
being 1: 00002728-n
body 3: 03607347-n
microorganism 1: 00740781-n
spiritual being 1: 05773239-n
Location+Solid
land 8: 08132366-n
MoneyRepresentation
financial obligation 1: 08222484-n
payment 2: 08147362-n
MoneyRepresentation+Artifact
medium of exchange 1: 08207032-n
money 1: 08132772-n
money 2: 08214427-n
money 3: 08214665-n
MoneyRepresentation+Group+Artifact
coinage 3: 08216671-n
MoneyRepresentation+Object+Artifact
coin 1: 08217024-n
currency 3: 08215253-n
MoneyRepresentation+Part+Artifact
amount of money 1: 08180701-n
Object
body 9: 05641227-n
complex 1: 03975160-n
stick 3: 02909904-n
Object+Animal
Equus caballus 1: 01691640-n
animal 1: 00008030-n
aquatic vertebrate 1: 00855637-n
arthropod 1: 01126858-n
bird 1: 00884285-n
canid 1: 01421448-n
carnivore 2: 01413653-n
chordate 1: 00849436-n
craniate 1: 00854210-n
dog 1: 01422174-n
equid 1: 01691356-n
eutherian 1: 01237932-n
fish 2: 01816356-n
hoofed mammal 1: 01688143-n
insect 1: 01491542-n
invertebrate 1: 01254383-n
larva 1: 01633257-n

mammal 1: 01213903-n
mollusc 1: 01286451-n
odd-toed ungulate 1: 01690543-n
offspring 1: 00736689-n
reptile 1: 01033306-n

Object+Artifact
artefact 1: 00011607-n
book 1: 02174965-n
construction 4: 02034531-n
flat solid 1: 03056705-n
pole 1: 02908961-n
rod 3: 02909423-n

Object+Human
European 1: 05873418-n
acquaintance 2: 05918609-n
adherent 1: 06048864-n
adult 2: 05839075-n
adult female 1: 06434591-n
adult male 1: 06193747-n
advocate 1: 05923094-n
artist 1: 05939406-n
assistant 1: 05940574-n
athlete 1: 05942710-n
boy 3: 06192735-n
caller 1: 05981698-n
child 1: 05996700-n
child 2: 05997221-n
communicator 1: 05842570-n
compeer 1: 05852391-n
connection 6: 06015983-n
contestant 1: 05843454-n
creator 1: 05844200-n
denizen 1: 05848227-n
expert 1: 05846273-n
family 6: 06163682-n
female 2: 05847495-n
follower 1: 06093600-n
friend 3: 06102108-n
homo 1: 01779125-n
human 1: 00004865-n
intellect 3: 05849094-n
leader 2: 05850058-n
life 6: 06178692-n
male 2: 05850734-n
man 5: 06194712-n
man 7: 06195173-n
native 1: 05848758-n
offspring 2: 06233328-n
relation 3: 06163124-n
religionist 1: 05853722-n
ruler 2: 06313765-n
unfortunate 1: 05855160-n

Object+Natural
Earth 1: 05696519-n
celestial body 1: 05698341-n
inanimate object 1: 00009469-n
natural object 1: 00009919-n

Object+Plant
bush 4: 07998630-n
graminaceous plant 1: 07072915-n
tree 1: 07991027-n

Occupation+Group+Human
business 8: 05155150-n
company 4: 05223147-n
company 6: 05232180-n

Occupation+Object+Human
Dr. 1: 06050986-n
artificer 2: 06026990-n
author 2: 06438760-n
chair 4: 06279934-n
chief 2: 06127722-n
employee 1: 06069879-n
entertainer 1: 05845591-n
functionary 1: 06232382-n
health care provider 1: 06128804-n
instrumentalist 1: 06219943-n
man 8: 06337508-n
medical man 1: 06203256-n
party 5: 06248866-n
performer 1: 06256875-n
president 1: 06279283-n
president 2: 06279719-n
professional 2: 06285396-n
skilled worker 1: 06349626-n
soldier 2: 06357018-n
worker 2: 05856677-n
Part
amount 1: 00018966-n
atom 1: 08803169-n
atom 2: 08803320-n
bound 2: 05383364-n
component 1: 02334827-n
division 4: 03973162-n
group 3: 08804621-n
part 10: 05650477-n
part 12: 08450839-n
part 3: 02855539-n
section 2: 02880516-n
unit 8: 08451350-n

Part+Human
department 1: 05189859-n

Part+Liquid+Living
body fluid 1: 03725816-n

Part+Living
anatomical structure 1: 03612911-n
body part 1: 03610098-n
cell 1: 00003711-n
contractile organ 1: 03645654-n
muscle 3: 03645458-n
organ 4: 03650737-n

Part+Object+Living
bone 2: 03634323-n

Part+Object+Plant
fruit 3: 08017859-n

Part+Plant
plant organ 1: 07977350-n
plant part 1: 07976849-n

Part+Solid
end 7: 05412066-n
end 8: 05412182-n

end 9: 05412624-n
section 9: 05652971-n
Part+Solid+Artifact
city 3: 05397774-n
piece of paper 1: 04141240-n
slip 9: 03141951-n
Part+Solid+Living
membrane 2: 03740823-n
tissue 1: 03632471-n
Part+Solid+Natural
earth 4: 08919214-n
Part+Solid+Plant
wood 4: 09057553-n
Part+Substance
layer 2: 02707655-n
Part+Substance+Living
body substance 1: 03631546-n
hormone 1: 03729776-n
secretion 1: 03728455-n
Part+Substance+Plant
foliage 2: 08032472-n
plant material 1: 09008290-n
Place
cosmos 2: 05655960-n
country 3: 05400698-n
course 4: 02955611-n
home 4: 05372409-n
line 21: 05432072-n
location 1: 00014314-n
municipality 2: 05447262-n
part 9: 05449837-n
place 10: 05444846-n
place 13: 05469653-n
point 12: 05443777-n
work 3: 01962095-n
Place+Artifact
city 2: 05390395-n
way 4: 02031514-n
Place+Part
administrative district 1: 05373867-n
area 1: 02075853-n
area 5: 05376564-n
district 1: 05404435-n
enclosure 2: 02472938-n
extremity 3: 05413816-n
gap 4: 05661636-n
geographic area 1: 05417924-n
opening 4: 02028879-n
province 1: 05463659-n
region 3: 05450515-n
side 1: 02487333-n
surface 1: 02486678-n
surface 4: 05467731-n
Place+Part+Artifact
excavation 3: 02480168-n
Place+Part+Liquid+Natural
body of water 1: 05715416-n
Place+Part+Natural
geographic point 1: 05420170-n
interstice 2: 03614829-n
Place+Part+Solid
athletic field 1: 05415062-n
face 12: 05382030-n
field 11: 05414707-n
layer 3: 05430251-n
parcel 4: 05472252-n
space 7: 05462485-n
Place+Part+Solid+Natural
dry land 1: 05720524-n
Place+Solid
location 4: 03531499-n
place 7: 05384109-n
Place+Solid+Artifact
road 2: 03001757-n
Place+Solid+Natural
depression 4: 05657514-n
elevation 6: 05657252-n
Place+Substance+Natural
formation 5: 05656341-n
Plant
fungus 1: 07910410-n
grass 2: 07073185-n
herb 2: 07169764-n
ligneous plant 1: 07990292-n
tracheophyte 1: 07974178-n
Representation
indication 1: 04430266-n
medium 3: 04140264-n
Representation+Artifact
meter reading 2: 03944736-n
sign 3: 04425761-n
song 3: 04567799-n
symbol 2: 04434881-n
Representation+Object+Artifact
biography 1: 04268429-n
calling card 1: 04337362-n
sign 4: 04427279-n
Representation+Part
section 4: 04213050-n
Representation+Solid+Artifact
card 6: 04263357-n
material 4: 04338410-n
Software+Artifact
computer program 1: 04297609-n
database 1: 04339764-n
list 1: 04248202-n
software 1: 04296594-n
Solid
fiber 3: 08932374-n
metal 1: 08807415-n
powder 2: 09012321-n
solid 3: 09033134-n
Solid+Artifact
paper 6: 08996165-n
thread 1: 02361568-n
Solid+Living
protein 1: 08849625-n
Solid+Natural
mineral 1: 08983367-n
rock 4: 05637686-n

rock 5: 08827122-n
Substance
agent 5: 08879673-n
alloy 2: 08783498-n
chemical compound 1: 08907331-n
chemical element 1: 08805286-n
coloring material 1: 09003076-n
drug 1: 02003723-n
element 7: 08918157-n
material 5: 08781633-n
matter 1: 00010368-n
mixture 5: 08783090-n
pigment 1: 09006729-n
poison 2: 09028514-n
salt 5: 09018436-n
Substance+Living
fat 3: 08930612-n
neoplasm 1: 08647560-n
Substance+Natural
deposit 4: 05659254-n
organic compound 1: 08849147-n
Vehicle+Artifact
conveyance 3: 01991412-n
Vehicle+Object+Artifact
aircraft 1: 02051671-n
auto 1: 02242147-n
automotive vehicle 1: 02799224-n
boat 1: 02167572-n
craft 2: 03235595-n
ship 1: 03061180-n
vehicle 1: 03233330-n

2ndOrderEntities
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
SituationType
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
SituationType
continue 7: 01517254-v
leave 4: 00079704-v
thing 11: 08533938-n
SituationType+Condition
hold 26: 01515519-v
SituationType+Experience+Mental
desire 4: 01040073-v
experience 6: 01008772-v
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Dynamic
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Dynamic
affair 1: 03869121-n
alter 2: 00071241-v
change 11: 00064108-v
come about 1: 00204516-v
passage 1: 00114479-n
Dynamic+Agentive
act 12: 01341700-v
carry out 4: 01448761-v
do 6: 00980842-v
Dynamic+Agentive+Communication
convey 1: 00522332-v
evince 1: 00531321-v
express 5: 00529407-v
give information 1: 00467082-v
mouth 6: 00530290-v
say 8: 00569629-v
Dynamic+Agentive+Communication+Social
cozen 3: 01456537-v
Dynamic+Agentive+Condition
development 1: 00139142-n
Dynamic+Agentive+Condition+Purpose
deed 1: 00020244-n
improvement 1: 00138272-n
Dynamic+Agentive+Condition+Purpose+Social
aid 1: 00383106-n
aid 2: 00664219-n
therapy 1: 00385186-n
Dynamic+Agentive+Existence+Purpose+Communication+Social
art 1: 00518008-n
Dynamic+Agentive+Experience+Physical
look 8: 01216027-v
Dynamic+Agentive+Location
conduct 5: 01141779-v
Dynamic+Agentive+Mental
act 2: 03885466-n
basic cognitive process 1: 03885854-n
Dynamic+Agentive+Mental+Purpose
arrange 2: 00416049-v
categorization 2: 03900455-n
cerebration 1: 03918967-n
higher cognitive process 1: 03918844-n
Dynamic+Agentive+Physical+Condition
clean 2: 00023287-v
clean 4: 00106393-v
clean 5: 00109110-v
Dynamic+Agentive+Physical+Condition+Purpose+Social
medical aid 1: 00384138-n
Dynamic+Agentive+Physical+Location
meeting 1: 00069655-n
Dynamic+Agentive+Physical+Location+Manner
foot 8: 01084973-v
Dynamic+Agentive+Physical+Location+Purpose
travel 2: 00166345-n
Dynamic+Agentive+Physical+Location+Purpose+Usage
eat 3: 00663538-v
Dynamic+Agentive+Physical+Purpose
clean 7: 00881979-v
sex 1: 00469903-n
Dynamic+Agentive+Physical+Purpose+Social
athletics 1: 00240760-n
dance 1: 00299543-n
Dynamic+Agentive+Purpose
activity 1: 00228990-n
carrying into action 1: 00055898-n
exert effort 1: 01366212-v
Dynamic+Agentive+Purpose+Communication+Social
language 5: 04598615-n
Dynamic+Agentive+Purpose+Possession+Social
exchange for money 1: 01277199-v
Dynamic+Agentive+Purpose+Social
action 2: 00527228-n
compete 1: 00605050-v
duty 1: 00398775-n
governance 1: 00622561-n
group action 1: 00597858-n
penalization 1: 00639819-n
play 21: 00605818-v
Dynamic+Agentive+Quantity
accumulate 2: 00796914-v
Dynamic+Agentive+Social
act together 2: 01346535-v
function 1: 00399406-n
Dynamic+Cause
act 1: 00016649-n
action 1: 00021098-n
allow 6: 01371393-v
alter 3: 00072540-v

alteration 3: 04697176-n
change of state 1: 00113334-n
Dynamic+Cause+Location
displace 3: 01055491-v
Dynamic+Cause+Physical
cover 16: 00763269-v
Dynamic+Cause+Physical+Location
cause to spread 1: 00792958-v
impel 1: 00869132-v
Dynamic+Cause+Physical+Location+Manner
push 1: 00064101-n
Dynamic+Cause+Purpose
means 1: 00096919-n
Dynamic+Cause+Purpose+Possession
cater 2: 00671827-v
Dynamic+Cause+Quantity
increase 6: 00091455-v
Dynamic+Cause+Time
pass 39: 01531792-v
Dynamic+Condition
ameliorate 2: 00123997-v
decline 5: 00122638-v
flush 4: 08682700-n
Dynamic+Experience
experience 7: 01203891-v
experience 8: 01204902-v
find 3: 00307705-v
reality 1: 03940989-n
Dynamic+Experience+Mental
cognition 1: 00012878-n
desire 2: 04788545-n
disposition 2: 03287725-n
disposition 4: 04113320-n
disturbance 7: 08693431-n
emotion 1: 04785784-n
feeling 1: 00013522-n
humor 3: 04827440-n
pleasance 1: 04792478-n
Dynamic+Experience+Mental+Existence
process 4: 03885684-n
Dynamic+Experience+Physical
feel 12: 01202814-v
Dynamic+Location
change position 1: 01043075-v
come down 3: 01122509-v
go 14: 01046072-v
travel 5: 01049627-v
turn 22: 01086483-v
Dynamic+Location+Manner
ride 8: 01114042-v
Dynamic+Phenomenal
action 7: 08239425-n
bad luck 1: 04701573-n
chance 3: 06467144-n
consequence 3: 06465491-n
natural phenomenon 1: 06464347-n
Dynamic+Phenomenal+Condition
symptom 2: 08671032-n
Dynamic+Phenomenal+Experience+Physical
phenomenon 1: 00019295-n
Dynamic+Phenomenal+Physical
atmospheric phenomenon 1: 06472551-n
biological process 1: 08258903-n
light 12: 06502153-n
physical phenomenon 1: 06467898-n
wind 7: 06529752-n
Dynamic+Phenomenal+Physical+Condition
growth 4: 08647140-n
Dynamic+Phenomenal+Physical+Location
come down 4: 01558020-v
Dynamic+Physical+Location
accumulate 3: 01311458-v
change of position 1: 00186555-n
divide 5: 01161526-v
locomotion 1: 00159178-n
motion 5: 04704743-n
Dynamic+Physical+Location+Manner
actuation 1: 00058021-n
Dynamic+Physical+Location+Purpose
journey 1: 00172823-n
Dynamic+Possession
acquire 3: 01261345-v
acquiring 1: 00041613-n
have 15: 01260836-v
lose 7: 01301277-v
Dynamic+Quantity
change magnitude 1: 00101800-v
decrease 5: 00090574-v
increase 7: 00093597-v
Dynamic+Stimulating
cause to be heard 1: 01241976-v
cause to be perceived 1: 01212141-v
Dynamic+Stimulating+Experience
trouble 3: 04692813-n
Dynamic+Stimulating+Experience+Mental
affect 5: 01007544-v
arouse 5: 01003070-v
excite 2: 01004175-v
Dynamic+Stimulating+Experience+Physical
perception 2: 03890199-n
sensation 1: 03892008-n
Dynamic+Stimulating+Experience+Physical+Communication
cause to appear 1: 01219939-v
Dynamic+Stimulating+Physical
emit 2: 00554586-v
%%%%%%%%%%%%%%%%%%%%%%%%%%%

BoundedEvent
become 1: 00089026-v
cease 2: 00211850-v
change state 1: 00086015-v
event 1: 00016459-n
happening 1: 04690182-n
BoundedEvent+Agentive
complete 2: 00285198-v
error 1: 00038929-n
failure 1: 00035229-n
let 4: 00433082-v

nonaccomplishment 1: 00035066-n
BoundedEvent+Agentive+Existence
creation 2: 00505014-n
BoundedEvent+Agentive+Existence+Purpose+Communication
enter 1: 00563886-v
BoundedEvent+Agentive+Experience+Condition+Purpose
examine 4: 01226339-v
BoundedEvent+Agentive+Mental
abandon 3: 00345074-v
ascertain 3: 00517007-v
call back 1: 00341396-v
BoundedEvent+Agentive+Mental+Communication
admit defeat 1: 00611702-v
BoundedEvent+Agentive+Mental+Existence++Purpose
devise 3: 00396499-v
BoundedEvent+Agentive+Mental+Existence+Purpose+Communication
account 13: 01289475-v
BoundedEvent+Agentive+Mental+Purpose
analyse 3: 00362566-v
cerebrate 1: 00354465-v
choice 1: 00091731-n
choose 1: 00379073-v
decide 1: 00392710-v
determine 2: 00393722-v
differentiate 4: 00365740-v
form an opinion of 1: 00376571-v
identify 2: 00348034-v
BoundedEvent+Agentive+Mental+Purpose+Communication
affirm 1: 00374169-v
BoundedEvent+Agentive+Mental+Purpose+Social
form a resolution about 1: 00392562-v
BoundedEvent+Agentive+Physical+Condition
carve up 1: 01396914-v
cleaning 1: 00139539-n
BoundedEvent+Agentive+Physical+Existence
create from raw material 1: 00945714-v
kill 1: 00124269-n
BoundedEvent+Agentive+Physical+Existence+Communication
describe 1: 00366972-v
represent 3: 00556972-v
BoundedEvent+Agentive+Physical+Existence+Condition
conserve 2: 01268422-v
BoundedEvent+Agentive+Physical+Existence+Purpose
make 15: 00929175-v
BoundedEvent+Agentive+Physical+Existence+Purpose+Communication
interpret 5: 00966090-v
BoundedEvent+Agentive+Physical+Location
bring 8: 01188762-v
cut 32: 00894185-v
BoundedEvent+Agentive+Physical+Location+Possession
bring 2: 00823804-v
bring 3: 00824200-v
BoundedEvent+Agentive+Physical+Location+Purpose
direct 10: 01100714-v
maneuver 3: 00323663-n
BoundedEvent+Agentive+Physical+Location+Purpose+Manner
blow 2: 00647048-n
BoundedEvent+Agentive+Physical+Location+Purpose+Possession
get rid of 2: 01267839-v
BoundedEvent+Agentive+Physical+Location+Purpose+Social+Manner
stroke 3: 00329906-n
BoundedEvent+Agentive+Physical+Purpose+Communication
sign 3: 04425761-n
sign 6: 04479492-n
BoundedEvent+Agentive+Physical+Purpose+Social
assail 1: 00633037-v
BoundedEvent+Agentive+Possession
give 16: 01254390-v
BoundedEvent+Agentive+Purpose
accomplishment 1: 00019847-n
assay 3: 01432563-v
operation 3: 00338477-n
BoundedEvent+Agentive+Purpose+Communication
ask 1: 00422854-v
declare 5: 00570287-v
explain 2: 00528672-v
BoundedEvent+Agentive+Purpose+Communication+Social
allow 3: 00451248-v
asking 1: 04638292-n
character 3: 04001822-n
order 6: 04629714-n
party 1: 04769704-n
party 2: 05255204-n
performance 4: 04487114-n
show 1: 00297544-n
show 3: 04326789-n
speech act 1: 04625000-n
statement 4: 04388724-n
BoundedEvent+Agentive+Purpose+Communication+Social+Manner
declaration 2: 04390828-n
BoundedEvent+Agentive+Purpose+Communication+Usage+Manner
rhetorical device 1: 04590378-n
BoundedEvent+Agentive+Purpose+Possession
gift 4: 01255335-v
transfer 12: 01266189-v
BoundedEvent+Agentive+Purpose+Possession+Social

make a payment 1: 01281885-v
BoundedEvent+Agentive+Purpose+Social
appoint 3: 01401683-v
attack 5: 00540241-n
battle 2: 00527805-n
check 28: 01421427-v
chore 1: 00398968-n
competition 3: 04771851-n
game 1: 00254052-n
operation 6: 00528736-n
war 1: 00540597-n
BoundedEvent+Agentive+Purpose+Usage
apply 4: 00658243-v
BoundedEvent+Agentive+Quantity
add 1: 00110396-v
decrease 6: 00262983-v
BoundedEvent+Agentive+Social
play 24: 00652908-v
project 2: 00442844-n
BoundedEvent+Cause
break 23: 00218979-v
bring 1: 00078946-v
cause 6: 00432532-v
cause to have 1: 01317872-v
cease 3: 01515268-v
change 1: 00108829-n
conclusion 2: 00119310-n
keep 12: 01387332-v
leave 6: 00291924-v
BoundedEvent+Cause+Condition
arrange 4: 00842219-v
bring to a close 1: 00402474-v
cause 7: 00941367-v
fail to keep 1: 01301401-v
BoundedEvent+Cause+Condition+Possession
fail to profit 1: 01302104-v
BoundedEvent+Cause+Existence
bring to an end 1: 00213455-v
production 1: 00507790-n
BoundedEvent+Cause+Experience+Physical
cause to feel unwell 1: 00040824-v
BoundedEvent+Cause+Physical
fasten 3: 00768642-v
forge 6: 00949570-v
form 12: 00083270-v
leave a mark on 1: 00297919-v
BoundedEvent+Cause+Physical+ +Location
collect 2: 00794237-v
BoundedEvent+Cause+Physical+Condition
adorn 2: 00959417-v
break 19: 00154558-v
break 21: 00201902-v
break 31: 00787971-v
injure 1: 00043545-v
BoundedEvent+Cause+Physical+Existence
create 1: 00926188-v
create 2: 00926361-v
create again 1: 00928226-v
BoundedEvent+Cause+Physical+Existence+
kill 5: 00758542-v
BoundedEvent+Cause+Physical+Location
close 5: 00772512-v
disunite 1: 00897572-v
hit 15: 00806352-v
lay 3: 00859635-v
BoundedEvent+Cause+Physical+Location+Manner
project through the air 1: 00867132-v
cause to move by striking 1: 00809580-v
BoundedEvent+Cause+Physical+Location+Possession
furnish 1: 01323715-v
BoundedEvent+Cause+Physical+Quantity
change of magnitude 1: 00196939-n
decrease 1: 00197092-n
increase 1: 00204508-n
BoundedEvent+Condition+Possession
loss 1: 00036401-n
BoundedEvent+Existence
constitution 1: 00134247-n
BoundedEvent+Experience+Existence+Time
life 13: 09084835-n
BoundedEvent+Experience+Mental
discover 5: 00937054-v
BoundedEvent+Experience+Time
night 5: 09100842-n
BoundedEvent+Location
arrive 1: 01144761-v
come 6: 01054590-v
come in 5: 01152122-v
depart 1: 01054314-v
go away 3: 01147140-v
go by 3: 01172741-v
BoundedEvent+Mental
bump into 2: 01280035-v
BoundedEvent+Phenomenal+Experience+Quantity+Time
dark 5: 09100431-n
BoundedEvent+Physical
change integrity 1: 00081466-v
connect 4: 00778333-v
BoundedEvent+Physical+Condition
break 20: 00201526-v
break into fragments 1: 00203548-v
break into parts 1: 00237247-v
BoundedEvent+Physical+Existence
decease 2: 00216283-v
BoundedEvent+Physical+Location
attach 3: 00743265-v
bring 5: 00827521-v
change of location 1: 00157028-n
collide with 1: 00704074-v
fill 5: 00268884-v
remove 2: 00104355-v
touch 18: 00686113-v
BoundedEvent+Physical+Location+Manner
stroke 2: 00318118-n
BoundedEvent+Physical+Location+Possession
get hold of 2: 00691086-v

BoundedEvent+Quantity
increase 3: 04725113-n
BoundedEvent+Quantity+Purpose+Time
day 5: 09094193-n
BoundedEvent+Quantity+Purpose+Usage+Time
time 9: 09171650-n
BoundedEvent+Quantity+Social+Time
day 3: 09081414-n
BoundedEvent+Quantity+Time
amount of time 1: 09065837-n
calendar day 1: 09094027-n
calendar month 1: 09131680-n
day 2: 09071807-n
day 4: 09092722-n
instant 1: 09157756-n
time 5: 09071447-n
twelvemonth 1: 09127492-n
year 2: 09125664-n
year 4: 09127774-n
BoundedEvent+Stimulating+Experience+Communication
express indirectly 1: 00469225-v
BoundedEvent+Stimulating+Physical
sound 5: 04731716-n
vocalization 1: 04599795-n
BoundedEvent+Stimulating+Purpose+Communication
demonstrate 1: 00373148-v
BoundedEvent+Stimulating+Purpose+Social
composition 8: 04561287-n
song 3: 04567799-n
BoundedEvent+Time
day 6: 09098948-n
day 7: 09130776-n
day 8: 09130983-n
night 4: 09100717-n
time 4: 04704458-n
BoundedEvent+Usage
break 26: 00258338-v
%%%%%%%%%%%%%%%%%%%%%%%%%%

UnboundedEvent
continue 2: 00210630-v
process 6: 08239006-n
UnboundedEvent+Agentive+Communication
communicate 1: 00416793-v
speak 2: 00542186-v
UnboundedEvent+Agentive+Communication+Manner
expressive style 1: 04575747-n
UnboundedEvent+Agentive+Condition+Purpose+Social
medical science 1: 04053427-n
UnboundedEvent+Agentive+Existence+Purpose+Communication
communicate by writing 1: 00559904-v
UnboundedEvent+Agentive+Mental
remember 2: 00342479-v
remember 3: 00343621-v
UnboundedEvent+Agentive+Mental+Purpose
abstract thought 1: 03919704-n
UnboundedEvent+Agentive+Mental+Purpose+Communication+Social
argumentation 1: 03920287-n
UnboundedEvent+Agentive+Physical+Condition+Purpose+Social
care for 1: 00048767-v
UnboundedEvent+Agentive+Physical+Manner
neaten 1: 00026120-v
UnboundedEvent+Agentive+Physical+Purpose+Manner
processing 1: 08300433-n
UnboundedEvent+Agentive+Physical+Social
fight 5: 00615347-v
UnboundedEvent+Agentive+Possession+Social
business 3: 00606634-n
UnboundedEvent+Agentive+Purpose+Communication+Social
communicating 1: 04138929-n
UnboundedEvent+Agentive+Purpose+Social
amusement 1: 00295035-n
biological science 1: 04052506-n
branch of knowledge 1: 04035790-n
business 2: 00341191-n
care for 4: 01378917-v
class 1: 00492074-n
command 10: 01381843-v
diversion 2: 00238878-n
head 28: 01381333-v
life science 1: 04052323-n
music 1: 00313161-n
natural philosophy 1: 04066626-n
natural science 1: 04037783-n
science 3: 04037371-n
social control 1: 00621770-n
work 1: 00337364-n
UnboundedEvent+Agentive+Social+Manner
act 7: 00007021-v
UnboundedEvent+Cause+Condition+Social
aid 6: 01442355-v
back up 4: 01446559-v
UnboundedEvent+Cause+Experience+Physical
cause pain 1: 00040663-v
UnboundedEvent+Condition
development 6: 08283435-n
UnboundedEvent+Experience
life 3: 03941565-n
UnboundedEvent+Experience+Existence
life 8: 08543710-n
UnboundedEvent+Experience+Time
time 1: 00014882-n
UnboundedEvent+Manner
pattern 1: 00230674-n
UnboundedEvent+Mental+Purpose+Social
science 2: 04037192-n
UnboundedEvent+Phenomenal+Physical

reaction 2: 00478685-n
UnboundedEvent+Physical
activity 4: 08274118-n
UnboundedEvent+Physical+Location+Purpose+Usage
consume 2: 00656714-v
UnboundedEvent+Physical+Purpose+Communication+Social
music 4: 04552184-n
UnboundedEvent+Social+Manner
behavior 3: 03433579-n
%%%%%%%%%%%%%%%%%%%%%%%%%%%

Static

%%%%%%%%%%%%%%%%%%%%%%%%%%%
Static
be 4: 01472320-v
continue 1: 00068138-v
position 12: 08522029-n
state 1: 00015437-n
thing 6: 03966203-n
union 9: 08711637-n
Static+Agentive+Purpose
arrangement 4: 03898749-n
Static+Cause+Purpose
system 4: 03864615-n
Static+Cause+Quantity
measure 5: 03539714-n
Static+Condition+Social
accord 4: 08549511-n
dignity 3: 08719491-n
disorder 1: 08550427-n
Static+Existence
death 5: 08781169-n
Static+Manner
fashion 2: 03450012-n
Static+Mental
abstract 1: 03965572-n
Static+Mental+Location
place 3: 03837930-n
Static+Phenomenal+Condition
atmospheric condition 1: 06529389-n
Static+Quantity
batch 3: 08432825-n
definite quantity 1: 08310215-n
indefinite quantity 1: 08310433-n
number 2: 03553723-n
quantity 3: 03966324-n
small indefinite quantity 1: 08423016-n
Static+Quantity+Purpose+Usage+Social
unit 6: 08313335-n
Static+Social
berth 1: 00344376-n
employment 1: 00342842-n
natural state 1: 08530753-n
Static+Stimulating+Mental
motivation 1: 00013299-n
%%%%%%%%%%%%%%%%%%%%%%%%%%%

Property

attribute 1: 00017586-n
be 8: 01482115-v
character 2: 03963513-n
end 16: 01475351-v
nature 2: 03340632-n
Property 2: 03444246-n
quality 1: 03338771-n
thing 4: 03283615-n
trait 1: 03282629-n
Property+Agentive+Purpose+Possession+Social
sell 7: 01546360-v
Property+Cause+Modal
can 8: 01539155-v
Property+Condition
condition 4: 08520221-n
condition 5: 08520394-n
defect 3: 08738373-n
deficiency 2: 08731035-n
need 5: 00675532-v
need 6: 00675686-v
situation 4: 08522741-n
Property+Condition+Social
value 2: 03564110-n
worth 1: 03563866-n
Property+Existence
be 3: 01471536-v
be 6: 01477879-v
Property+Experience+Mental
cognize 1: 00333362-v
understand 1: 00330150-v
Property+Experience+Physical+Modal
sense 2: 03858744-n
Property+Mental
await 1: 00405636-v
believe 3: 00387631-v
consider 1: 00388394-v
psychological feature 1: 00012517-n
Property+Mental+Communication+Social
agree 2: 00452960-v
Property+Mental+Modal
faculty 1: 03857413-n
Property+Mental+Purpose
way 7: 03930651-n
Property+Modal
ability 1: 03601639-n
ability 2: 03841132-n
appear 6: 01217877-v
inability 2: 03854243-n
Property+Physical
form 1: 00014558-n
Property+Physical+Condition
be ill with 1: 00041140-v
disease 1: 08592183-n
disorder 2: 08586618-n
harm 3: 08665752-n
health problem 1: 08586350-n
illness 1: 08587853-n

physiological state 1: 08577911-n
plant disease 1: 08658681-n
Property+Physical+Location+Possession
carry 27: 01537537-v
Property+Physical+Manner
structure 2: 03451157-n
style 6: 03961040-n
Property+Physical+Quantity
magnitude 1: 03539122-n
Property+Purpose+Modal
accomplishment 2: 03849803-n
Property+Purpose+Social
agency 3: 08565692-n
Property+Quantity
number 10: 08317731-n
number 5: 04231864-n
Property+Social+Modal
play 16: 08569341-n
potency 2: 03596179-n
Property+Stimulating+Physical
appearance 4: 03314728-n
cast 7: 03316776-n
color 2: 03463765-n
form 6: 04003083-n
visual property 1: 03460270-n
Property+Time
time 6: 09077332-n
%%%%%%%%%%%%%%%%%%%%%%%%%%%

Relation

agree 5: 01503041-v
connectedness 1: 08440487-n
degree 1: 03540591-n
relation 1: 00017862-n
relationship 1: 08436181-n
Relation+Agentive+Purpose+Communication
intend 4: 00537777-v
Relation+Communication
be about 2: 01513147-v
Relation+Condition+Social
degree 7: 08535290-n
position 13: 08534455-n
Relation+Location
be 9: 01501697-v
course 8: 05666985-n
degree 6: 08531278-n
direction 7: 05477069-n
go 25: 01518088-v
space 1: 00015245-n
spacing 1: 03535737-n
stay in one place 1: 01492762-v
Relation+Physical+Location
adjoin 1: 00685874-v
aim 4: 05477280-n
blank space 1: 04211782-n
course 7: 05477560-n
direction 8: 08463109-n
distance 1: 03536009-n
elbow room 1: 08434357-n
path 3: 05441398-n
spatial property 1: 03524985-n
spatial relation 1: 08462976-n
Relation+Physical+Quantity
magnitude relation 1: 08454813-n
ratio 1: 08457189-n
Relation+Possession
have 12: 01256853-v
have 13: 01257491-v
hold on to 2: 01256282-v
Relation+Quantity
be 10: 01506899-v
Relation+Social
family relationship 1: 08453309-n
rank 3: 08717824-n
relationship 3: 08523567-n
relationship 4: 08523811-n
social relation 1: 00018392-n

3rdOrderEntity

3rdOrderEntity+Cause+Mental+Purpose
plan 3: 03985547-n
plan of action 1: 03987224-n
procedure 3: 00566905-n
3rdOrderEntity+Cause+Mental+Purpose+Communication+Social
policy 3: 04349399-n
3rdOrderEntity+Cause+Mental+Purpose+Social
play 7: 00324581-n
3rdOrderEntity+Experience+Mental
attitude 3: 04111788-n
faith 2: 04011318-n
know-how 1: 03841532-n
3rdOrderEntity+Mental
belief 2: 04008826-n
category 1: 03957148-n
cognitive content 1: 03940357-n
concept 1: 03954891-n
data point 1: 03944568-n
doctrine 1: 04009596-n
evidence 1: 03948538-n
idea 2: 03953834-n
info 1: 04337839-n
information 1: 03944302-n
issue 4: 03943820-n
knowledge base 1: 04036935-n
opening 7: 03930751-n
opinion 2: 04010732-n
structure 4: 03898550-n
subject 5: 04314223-n
theory 3: 04033925-n
thing 8: 04389685-n
3rdOrderEntity+Mental+Communication+Usage
message 2: 04313427-n
3rdOrderEntity+Mental+Purpose+Communication+Social
communication 1: 00018599-n
3rdOrderEntity+Mental+Purpose+Manner
method 2: 03863261-n
3rdOrderEntity+Mental+Social
right 4: 03586387-n
3rdOrderEntity+Stimulating+Mental
life 5: 05633277-n
3rdOrderEntity+Stimulating+Mental+Purpose
aim 2: 04029556-n
aim 3: 04030116-n

Appendix III: Top Concept Cluster Combinations for Base Concepts

```
1    3rdOrderEntity;Cause;Mental;Purpose;Communication;Social
1    3rdOrderEntity;Cause;Mental;Purpose;Social;Recreation
1    3rdOrderEntity;Experience;Mental;cognition
1    3rdOrderEntity;Mental;information,cognition
1    3rdOrderEntity;Mental;Communication;Usage;information
1    3rdOrderEntity;Mental;Purpose;Communication;Social;cognition
1    3rdOrderEntity;Mental;Purpose;Manner
1    3rdOrderEntity;Mental;Social
1    3rdOrderEntity;Stimulating;Mental
2    3rdOrderEntity;Experience;Mental
2    3rdOrderEntity;Stimulating;Mental;Purpose
3    3rdOrderEntity;Cause;Mental;Purpose
3    3rdOrderEntity;Mental;information
7    3rdOrderEntity;Mental
7    3rdOrderEntity;Mental;cognition
```

1 BoundedEvent;Agentive;Existence
1 BoundedEvent;Agentive;Existence;Purpose;Communication
1 BoundedEvent;Agentive;Experience;Condition
1 BoundedEvent;Agentive;Mental;Communication
1 BoundedEvent;Agentive;Mental;Existence;Communication
1 BoundedEvent;Agentive;Mental;Existence;Purpose
1 BoundedEvent;Agentive;Mental;Purpose;cognition
1 BoundedEvent;Agentive;Mental;Purpose;Communication
1 BoundedEvent;Agentive;Mental;Purpose;Social
1 BoundedEvent;Agentive;Physical;Location;Purpose;Manner;conflict
1 BoundedEvent;Agentive;Physical;Location;Purpose;movement
1
BoundedEvent;Agentive;Physical;Location;Purpose;Social;Manner;Recreat
ion
1 BoundedEvent;Agentive;Physical;Purpose;Social;Fighting
1 BoundedEvent;Agentive;Purpose;Communication;Social;Manner
1 BoundedEvent;Agentive;Purpose;Communication;Usage;Manner
1 BoundedEvent;Agentive;Purpose;Social;Work
1 BoundedEvent;Agentive;Purpose;Usage
1 BoundedEvent;Agentive;Social;Games
1 BoundedEvent;Agentive;Social;Work
1 BoundedEvent;Cause;Condition;Possession
1 BoundedEvent;Cause;Experience;Physical
1 BoundedEvent;Cause;Physical;Location;Possession
1 BoundedEvent;Condition;Possession
1 BoundedEvent;Experience;Existence;Time
1 BoundedEvent;Experience;Mental
1 BoundedEvent;Experience;Time
1 BoundedEvent;Mental
1 BoundedEvent;Phenomenal;Experience;Quantity;Time
1 BoundedEvent;Physical;Existence
1 BoundedEvent;Physical;Location;Manner
1 BoundedEvent;Physical;Location;movement
1 BoundedEvent;Physical;Location;Possession
1 BoundedEvent;Quantity
1 BoundedEvent;Quantity;Purpose;Time
1 BoundedEvent;Quantity;Purpose;Usage;Time
1 BoundedEvent;Quantity;Social;Time;Work
1 BoundedEvent;Quantity;Time;Science
1 BoundedEvent;Quantity;Time;science
1 BoundedEvent;Stimulating;Experience;Communication
1 BoundedEvent;Stimulating;Purpose;Communication
1 BoundedEvent;Stimulating;Purpose;Social
1 BoundedEvent;Stimulating;Purpose;Social;Art
1 BoundedEvent;Usage
1 Dynamic;Agentive;Communication;Social;Behavior
1 Dynamic;Agentive;Condition
1 Dynamic;Agentive;Existence;Purpose;Communication;Social;Art
1 Dynamic;Agentive;Experience;Physical
1 Dynamic;Agentive;Location
1 Dynamic;Agentive;Location;Manner
1 Dynamic;Agentive;Mental;Purpose
1 Dynamic;Agentive;Physical;Condition;Chemistry
1 Dynamic;Agentive;Physical;Condition;Purpose;Social;Caring
1 Dynamic;Agentive;Physical;Location;movement
1 Dynamic;Agentive;Physical;Location;Purpose;movement
1 Dynamic;Agentive;Physical;Location;Purpose;Usage
1 Dynamic;Agentive;Physical;Purpose
1 Dynamic;Agentive;Physical;Purpose;Behavior
1 Dynamic;Agentive;Physical;Purpose;Social;Art
1 Dynamic;Agentive;Physical;Purpose;Social;Recreation

1 Dynamic;Agentive;Possession
1 Dynamic;Agentive;Purpose;Communication;Social
1 Dynamic;Agentive;Purpose;Social;Behavior
1 Dynamic;Agentive;Purpose;Social;conflict
1 Dynamic;Agentive;Purpose;Social;Management
1 Dynamic;Agentive;Purpose;Social;Recreation
1 Dynamic;Agentive;Purpose;Social;Work
1 Dynamic;Agentive;Quantity
1 Dynamic;Agentive;Social;Behavior
1 Dynamic;Agentive;Social;Work
1 Dynamic;Cause;Location
1 Dynamic;Cause;Physical
1 Dynamic;Cause;Physical;Location;Manner
1 Dynamic;Cause;Purpose;Possession
1 Dynamic;Cause;Quantity
1 Dynamic;Cause;Time
1 Dynamic;Experience;Mental;Existence
1 Dynamic;Experience;Physical
1 Dynamic;Location;Manner
1 Dynamic;Phenomenal;Condition
1 Dynamic;Phenomenal;Experience;Physical
1 Dynamic;Phenomenal;Physical;Condition
1 Dynamic;Phenomenal;Physical;Location;Wheather
1 Dynamic;Physical;Location;Manner;movement
1 Dynamic;Physical;Location;Purpose;movement
1 Dynamic;Quantity;Possession
1 Dynamic;Stimulating;Experience
1 Dynamic;Stimulating;Experience;Physical;Communication
1 Dynamic;Stimulating;Physical
1 SituationType
1 UnboundedEvent;Agentive;Communication;Manner
1 UnboundedEvent;Agentive;Condition;Purpose;Social;Science
1 UnboundedEvent;Agentive;Existence;Purpose;Communication
1 UnboundedEvent;Agentive;Mental;Purpose;cognition
1 UnboundedEvent;Agentive;Mental;Purpose;Communication;Social;cognition
1 UnboundedEvent;Agentive;Physical;Condition;Purpose;Social;Caring
1 UnboundedEvent;Agentive;Physical;Manner
1 UnboundedEvent;Agentive;Physical;Purpose;Manner
1 UnboundedEvent;Agentive;Physical;Social;Fighting
1 UnboundedEvent;Agentive;Possession;Social
1 UnboundedEvent;Agentive;Purpose;Communication;Social
1 UnboundedEvent;Agentive;Purpose;Social
1 UnboundedEvent;Agentive;Purpose;Social;Art
1 UnboundedEvent;Agentive;Purpose;Social;Education
1 UnboundedEvent;Agentive;Social;Manner;Behavior
1 UnboundedEvent;Cause;Experience;Physical
1 UnboundedEvent;Condition
1 UnboundedEvent;Experience
1 UnboundedEvent;Experience;Existence
1 UnboundedEvent;Experience;Time
1 UnboundedEvent;Manner
1 UnboundedEvent;Mental;Purpose;Social
1 UnboundedEvent;Phenomenal;Physical
1 UnboundedEvent;Physical
1 UnboundedEvent;Physical;Location;Purpose;Usage
1 UnboundedEvent;Physical;Purpose;Communication;Social;Art
1 UnboundedEvent;Social;Manner;Behavior
2 BoundedEvent;Agentive;Physical;Condition
2 BoundedEvent;Agentive;Physical;Purpose;Communication
2 BoundedEvent;Agentive;Purpose
2 BoundedEvent;Agentive;Purpose;Communication;Social;Recreation

2	BoundedEvent;Agentive;Purpose;Social;Management
2	BoundedEvent;Agentive;Purpose;Social;Recreation
2	BoundedEvent;Agentive;Quantity
2	BoundedEvent;Cause;Existence
2	BoundedEvent;Cause;Physical;Location;Manner
2	BoundedEvent;Existence
2	BoundedEvent;Physical
2	BoundedEvent;Stimulating;Physical
2	Dynamic;Agentive;Condition;Purpose
2	Dynamic;Agentive;Mental;cognition
2	Dynamic;Agentive;Physical;Condition
2	Dynamic;Agentive;Purpose
2	Dynamic;Agentive;Purpose;Social
2	Dynamic;Cause;Physical;Location
2	Dynamic;Cause;Purpose
2	Dynamic;Physical;Location;movement
2	Dynamic;Stimulating
2	Dynamic;Stimulating;Experience;Physical
2	SituationType;Experience;Mental
2	UnboundedEvent
2	UnboundedEvent;Agentive;Communication
2	UnboundedEvent;Agentive;Mental
2	UnboundedEvent;Agentive;Purpose;Social;Recreation
2	UnboundedEvent;Agentive;Purpose;Social;Work
2	UnboundedEvent;Cause;Condition;Social;Caring
3	BoundedEvent;Agentive;Physical;Existence
3	BoundedEvent;Agentive;Physical;Existence;Communication
3	BoundedEvent;Agentive;Physical;Location
3	BoundedEvent;Agentive;Physical;Location;Possession
3	BoundedEvent;Agentive;Purpose;Communication
3	BoundedEvent;Cause;Physical;Quantity
3	BoundedEvent;Physical;Condition
3	Dynamic;Agentive;Condition;Purpose;Social;Caring
3	Dynamic;Agentive;Mental;Purpose;cognition
3	Dynamic;Condition
3	Dynamic;Physical;Location
3	Dynamic;Quantity
3	Dynamic;Stimulating;Experience;Mental
3	SituationType;Cause
3	UnboundedEvent;Agentive;Purpose;Social;Management
4	BoundedEvent
4	BoundedEvent;Agentive;Mental
4	BoundedEvent;Agentive;Possession
4	BoundedEvent;Agentive;Purpose;Communication;Social;Art
4	BoundedEvent;Agentive;Purpose;Social;conflict
4	BoundedEvent;Cause;Condition
4	BoundedEvent;Cause;Physical;Condition
4	BoundedEvent;Cause;Physical;Existence
4	Dynamic;Agentive
4	Dynamic;Experience
4	Dynamic;Possession
5	BoundedEvent;Agentive;Purpose;Communication;Social
5	BoundedEvent;Cause;Physical
5	BoundedEvent;Cause;Physical;Location
5	BoundedEvent;Time
5	Dynamic
5	Dynamic;Location
5	Dynamic;Phenomenal
5	Dynamic;Phenomenal;Physical
6	BoundedEvent;Agentive
6	BoundedEvent;Location

6	**BoundedEvent;Physical;Location**
6	**Dynamic;Agentive;Communication**
6	**Dynamic;Cause**
6	**UnboundedEvent;Agentive;Purpose;Social;Science**
8	**BoundedEvent;Agentive;Mental;Purpose**
8	**BoundedEvent;Quantity;Time**
9	**BoundedEvent;Cause**
9	**Dynamic;Experience;Mental**

1	Static;Agentive;Purpose;cognition
1	Static;Cause;Purpose;behavior
1	Static;Cause;Quantity
1	Static;Condition;Social;Work
1	Static;Existence
1	Static;Manner;behavior
1	Static;Mental;cognition
1	Static;Mental;Location
1	Static;Phenomenal;Condition
1	Static;Quantity;Purpose;Usage;Social
1	Static;Social
1	Static;Stimulating;Mental
1	Property;Cause;Modal
1	Property;Experience;Physical;Modal
1	Property;Location;Possession
1	Property;Mental;Communication;Social
1	Property;Mental;Modal;cognition
1	Property;Mental;Purpose
1	Property;Physical
1	Property;Physical;Quantity
1	Property;Possession;Social
1	Property;Purpose;Modal
1	Property;Purpose;Social
1	Property;Time
1	Relation;Agentive;Purpose;Communication
1	Relation;Communication
1	Relation;Quantity
2	Static;Condition;Social
2	Static;Social;Work
2	Property;Condition;Social
2	Property;Existence
2	Property;Experience;Mental
2	Property;Physical;Manner
2	Property;Quantity
2	Property;Social;Modal
2	Relation;Condition;Social
2	Relation;Physical;Quantity
3	Property;Physical;Condition;health
3	Relation;Possession
4	Property;Mental
4	Property;Modal
5	Property;Physical;Condition
5	Property;Stimulating;Physical
5	Relation
5	Relation;Social
6	Static
6	Static;Quantity
7	Property;Condition
8	Relation;Location
9	Property
10	Relation;Physical;Location

1	1stOrderEntity
1	Building;Group;Artifact
1	Building;Object
1	Comestible;Group;Artifact
1	Comestible;Group;Plant
1	Comestible;Part
1	Comestible;Part;Solid
1	Comestible;Part;Solid;Natural
1	Comestible;Solid
1	Comestible;Solid;Animal
1	Container
1	Container;Object;Artifact
1	Container;Solid;Artifact
1	Covering
1	Covering;Artifact
1	Covering;Object;Natural
1	Covering;Part;Solid;Natural
1	Covering;Solid;Artifact
1	Function;Composition;Form;Origin
1	Function;Object;Artifact
1	Function;Part;Object;Artifact
1	Function;Solid;Natural
1	Furniture;Group;Artifact
1	Gas
1	Group;Living
1	Group;Plant
1	ImageRepresentation;Object
1	Instrument;Group
1	LanguageRepresentation;Group
1	Location;Solid
1	MoneyRepresentation
1	MoneyRepresentation;Group;Artifact
1	MoneyRepresentation;Part;Artifact
1	Part;Liquid;Living
1	Part;Object;Living
1	Part;Object;Plant
1	Part;Solid;Natural
1	Part;Solid;Plant
1	Part;Substance
1	Place;Part;Artifact
1	Place;Part;Liquid;Natural
1	Place;Part;Solid;Natural
1	Place;Solid;Artifact
1	Place;Substance;Natural
1	Representation;Part
1	Solid;Living
1	Vehicle;Artifact
2	Artifact
2	Building;Group;Object;Artifact
2	Building;Part;Object;Artifact
2	Comestible;Liquid
2	Comestible;Object;Plant
2	Container;Object
2	Container;Solid
2	Creature
2	ImageRepresentation
2	ImageRepresentation;Object;Artifact
2	Instrument;Group;Artifact
2	LanguageRepresentation;Part;Artifact
2	LanguageRepresentation;Solid;Artifact
2	MoneyRepresentation;Object;Artifact

2	**Occupation;Group;Human**
2	**Part;Plant**
2	**Part;Solid;Living**
2	**Part;Substance;Plant**
2	**Place;Part;Natural**
2	**Place;Solid**
2	**Place;Solid;Natural**
2	**Representation**
2	**Representation;Solid;Artifact**
2	**Solid;Artifact**
2	**Substance;Living**
2	**Substance;Natural**
3	**Comestible;Liquid;Artifact**
3	**Covering;Part;Solid;Living**
3	**Garment;Solid;Artifact**
3	**LanguageRepresentation;Object;Artifact**
3	**Object**
3	**Object;Plant**
3	**Part;Solid;Artifact**
3	**Part;Substance;Living**
3	**Representation;Object;Artifact**
3	**Solid;Natural**
4	**Comestible**
4	**Comestible;Substance**
4	**Function;Artifact**
4	**Function;Group;Human**
4	**ImageRepresentation;Artifact**
4	**MoneyRepresentation;Artifact**
4	**Object;Natural**
4	**Part;Solid**
4	**Representation;Artifact**
4	**Software;Artifact**
4	**Solid**
5	**Comestible;Artifact**
5	**Comestible;Solid;Artifact**
5	**Container;Part;Solid;Living**
5	**Furniture;Object;Artifact**
5	**Instrument;Artifact**
5	**Living**
5	**Plant**
6	**Liquid**
6	**Object;Artifact**
6	**Part;Living**
6	**Place;Part;Solid**
7	**Building;Object;Artifact**
7	**Group**
7	**LanguageRepresentation**
7	**Vehicle;Object;Artifact**
10	**Instrument;Object;Artifact**
12	**Part**
14	**Place**
14	**Place;Part**
15	**Substance**
19	**LanguageRepresentation;Artifact**
20	**Occupation;Object;Human**
22	**Object;Animal**
26	**Function**
38	**Group;Human**
42	**Object;Human**

Appendix IV EuroWordNet Optional Variant Information

Important Comments

The tables provided here reflect the situation of the current version 1.3 EWN database. When preparing import data, and you need to refer to a usage label or feature, use the string of the "Code" columns below. Do not use the string in the "Name" column. When preparing import data, and you need to refer to a value, use the string in the "Values" column. Do not use the numeric identifiers from the table below. For further information on the import/export syntax, please refer to the Polaris documentation.

Usage Labels

Language-independent Usage Labels

Name		**Code**	**Values**	
Date	1	date	Old-fashioned *archaic, out-of-date, obsolete*	1
Register	2	reg	Unusual *rare, infrequent*	1
			Usual *common, frequent*	2
			Formal *traditional, conventional, literary*	3
			Informal *familiar, unliterary, conversational*	4
			Humerous *comical*	5
			Poetic *literary*	6
			Vulgar *plebeian, rude, taboo*	7
			Slang *argot, used by certain social groups*	8
			Neologism *newly invented word*	9
			Burlesque *caricature, parody*	10
			Pejorative *negative, showing disapproval, uncomplementary*	11
			Positive *showing approval, complementary*	12
			Euphemistic *inexplicit, understatement*	13
			Ironic *sarcastic*	14
			Diminutive *small, little*	15
Sublanguage	3	sub	Scientific	1
			Technical	2
			Business	3
			Geography	4
			Medicine	5
			Computer	6
			Sports & Leisure	7
Origin	4	orig	Spanish	1
			German	2
			Latin	3
			French	4
			English	5
			Russian	6

Dutch Usage Labels

Name		Code	Values	
Dialect/Regional	1	dial	AZN	1
			Antilles	2

There are currently no usage labels for other languages.

Syntactic Features

Currently, all syntactic features are language-independent.

Name		Code	Values (1)	Parts-of-speech (2)
Gender	1	gender	masculine feminine neutral	n, v, a, b, p
Person	2	person	1st person singular 2nd person singular 3rd person singular 1st person plural 2nd person plural 3rd person plural polite singular polite plural	n, v, a, b, p
Number	3	number	singular plural dual	n, v, a, b, p
Tense	4	tense	...	n, v, a, b, p
Determiner	5	determiner	always never optional	n, v, a, b, p
Connotation	10	connotation	figurative non-figurative	n, v, a, b, p
Collective	101	collective	*	n
Countability	102	count	*	n
Portion	103	portion	*	n
Finite clause	104	fin_clause	*	n, v
Infinite clause	105	inf_clause	*	n, v
Nominal complement	106	nom_comp	*	n
Case	107	case	nom gen dat acc abl voc dual	n
Transitive	108	trans	*	v

Intransitive	109	intrans	*	v
Reflexive	110	reflexive	*	v
Middle formation	111	middle	*	v
Imperative form	112	imperative	*	v
Passive transformation	113	passive	*	v
Unaccusative	114	unacc	*	v
Unergative	115	unerg	*	v
Cognate object	116	cogn_obj	*	v
Empty object	117	empty_obj	*	v
Obligatory adverb	118	obl_adv	*	v
Obligatory negative polarity element	119	obl_neg_pol	*	v
Benefactive	120	benefact	*	v
Auxiliary for perfect tense	121	aux_perf	...	v
Status[(3)]	122	status	...	v
Prepositional object	123	prep_obj	...	v
Prepositional comitative	124	prep_comit	...	v
Prepositional object complement	125	prep_obj_comp	...	v
Prepositional copular verb	126	prep_cop	...	v
Locative	127	loc	...	v
Source	128	source	...	v
Target	129	target	...	v

1. In the "Values" column, if three periods appear instead of a list of values, it means that any text can be specified. If an asterisk (*) appears there instead, it means that the feature is a boolean value.
2. Part-of-speech codes are: n (noun), v (verb), a (adjective), b (adverb), p (proper noun)
3. Do not confuse this verb-specific feature field with the general 'Status' field.