VU VRIJE UNIVERSITEIT AMSTERDAM

VU Research Portal

The EuroWordNet Base Concepts and Top Ontology

Vossen, P.J.T.M.; Bloksma, L.; Rodriquez, H.

1998

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Vossen, P. J. T. M., Bloksma, L., & Rodriquez, H. (1998). *The EuroWordNet Base Concepts and Top Ontology*. Vrije Universiteit Amsterdam.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 03. Nov. 2026

The EuroWordNet Base Concepts and Top Ontology

Version 2, Final

January 22, 1998

Contributors:

Piek Vossen, Laura Bloksma, University of Amsterdam

Horacio Rodriguez, Politecnica de Catalunya (Barcelona)

Salvador Climent, University of Barcelona

Nicoletta Calzolari, Adriana Roventini, Francesca Bertagna, Antonietta Alonge, Istituto di Linguistica del CNR, Pisa

Wim Peters, University of Sheffield



Deliverable D017, D034, D036, WP5
EuroWordNet, LE2-4003

Identification number	LE-4003-D-017, D-034, D-036
Type	Document and Lingware
Title	The EuroWordNet Base Concepts and Top Ontology
Status	Final
Deliverable	D-017, D-034, D-036
Work Package	WP5
Task	n.a.
Period covered	June 1997 – December 1998
Date	January 22, 1998
Version	2
Number of pages	50
Authors	• Piek Vossen, Laura Bloksma, University of Amsterdam • Horacio Rodriguez, Politecnica de Catalunya (Barcelona) • Salvador Climent, University of Barcelona • Adriana Roventini, Francesca Bertagna, Antonietta Alonge, Istituto di Linguistica del CNR, Pisa • Wim Peters, University of Sheffield
WP/Task responsible	Amsterdam
Project contact point	Piek Vossen University of Amsterdam Spuistraat 134 1012 VB Amsterdam The Netherlands tel. +31 20 525 4669 fax. +31 20 525 4429 e-mail: Piek.Vossen@let.uva.nl
EC project officer	Ray Hudson
Status	Public
Actual distribution	Project Consortium, the EuroWordNet User Group, the Ad Hoc Ansii Committee for Ontology standards, the world via http://www.let.uva.nl/~ewn.
Suplementary notes	n.a.
Key words	Linguistic Resources, Ontologies, Multilingual Wordnets, Language Engineering

Abstract	This deliverable describes the EuroWordNet top ontology, the selection of the Base Concepts and the classification of these Base Concepts in terms of the top ontology.
Status of the abstract	Complete
Received on	
Recipient's catalogue number	

Executive Summary

In this document we describe the EuroWordNet Top Ontology and the selection of Base Concepts which have been classified using this ontology. In EuroWordNet we have chosen for a flexible design in which local wordnets can be built relatively independently as language-specific structures. For this purpose the word meanings in the local synsets are linked to an Inter-Lingual-Index or ILI. Furthermore, a strict separation is made between the language-specific wordnets and the language-independent modules. The separate construction of wordnets at different sites and the modular design of EuroWordNet also creates a disadvantage: the interpretation and coverage of the different wordnets may easily drift apart. Cross-lingual comparison and transfer of information across wordnets is only an option when the information in each wordnet is coded in a more-or-less compatible way. There are two specific compatibility issues at stake:

- the coverage of the vocabulary: the wordnets should globally deal with the same conceptual areas or domains
- the interpretation of the relations should be the same for all the sites.

To achieve maximal consensus, the wordnets are therefore developed top-down starting with a shared set of 1059 so-called Base Concepts which are classified using a common shared semantic framework. These Base Concepts are the most important meanings prevailing in the local wordnets. The Base Concepts represent the shared cores of the different wordnets, where we try to achieve a maximum of consensus and overlap. Still, the local wordnets can differ in the exact way in which the vocabulary is lexicalized around these Base Concepts.

To get to grips with these meanings, they have been classified by a Top Ontology, specifically designed for that purpose. The Top Ontology consists of 63 fundamental semantic distinctions used in various semantic theories and paradigms. Definitions and motivations are provided for each of these Top Concepts, and they have explicitly been defined in terms of hyponymy and opposition relations. The Top Ontology is linked to the ILI, thus providing some language-independent structuring of the ILI. The Top Concepts can be transferred via the equivalence relations of the ILI-records to the language-specific meanings, i.e.: they indirectly apply to all language-specific meanings related to the ILI-records that are classified in terms of the Top Concepts. Via the language-internal relations the Top Concept can be further inherited to all other related language-specific concepts. In this way the wordnets are built using a common framework, where the lexicalizations around these Base Concepts may differ from language to language.

This document describes in detail the selection of the Base Concepts and the Top Ontology that has been used to classify them. In the first section we further motivate the general top-down approach we have adopted. In section 3 we shortly describe the resources and methodologies of each site to clarify the different backgrounds and starting points. In section 4 we describe the technical procedure by which the set of common Base Concepts (BCs) has been established. Section 5 then describes a Top-Ontology of basic semantic distinctions, which has been developed to get to grips with these BCs. All BCs have been clustered in terms of these Top-Ontology Concepts providing a shared descriptive framework for the covered vocabulary In Appendix II, the classifications of the Base Concepts is given. Electronic versions can be down-loaded from http://www.let.uva.nl/~ewn.

Table of Contents

1 Introduction 6
2. The general approach for building the wordnets 8
3. Local Resources and Methods 10
4. The Base Concepts 12
4.1. Definition of the local BCs 12
4.2. Deriving a common set of Base Concepts 13
4.3. Tuning of the initial Common Base Concepts 15
5. Top-Ontology clustering of the Base Concepts 18
5.1. Starting points for the EuroWordNet Top-Ontology 19
5.2. The EuroWordNet Top-Ontology 21
5.2.1. Classification of 1st-Order-Entities 23
5.2.2. The classification of 2ndOrderEntities 26
6. Conclusions 31
References 32
Appendix I: Base Concepts Selected by four sites in EuroWordNet 33
Appendix II Top Ontology Classification of the Base Conceps 34
Appendix III: Top Concept Cluster Combinations for Base Concepts 45

1 Introduction

The aim of EuroWordNet (LE2-4003) is to build a multilingual database with wordnets for several languages. The wordnets are structured in the same way as the Princeton WordNet1.5 [Miller et al 1990] around the notion of a synset, which is a set of synonymous word meanings, and basic semantic relations such as hyponymy and meronymy between these synsets. In addition to the relations between the synsets of the separate languages there is also an equivalence relation for each synset to the closest concept from an Inter-Lingual-Index (ILI). The ILI contains all WordNet1.5 synsets but is extended with any other concept needed to establish precise equivalence relation across synsets. Via the equivalence relations with the ILI, it is possible to go from one wordnet to another wordnet, and compare synsets and their relations across languages.

The general approach of EuroWordNet is to build the wordnets **mainly** from existing resources. Each site in the project will be responsible for their language-specific wordnet using their tools and resources built up in previous national and international projects (Acquilex, Sift, Delis, Parole, Novell-ConceptNet, Van Dale, Bibliograf). This is not only more cost-effective but also gives us the possibility to combine and compare information from multiple independently-created resources. A comparison may tell us something about the vocabularies of the languages (typical lexicalization patterns) or about the consistency and quality of the resources (how much coherence is there across the resources, how rich are the resources compared to each other).

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 in the multilingual database. 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 expressed via the Inter-Lingual-Index (ILI). Language-specific synsets linked to the same ILI-record should thus be equivalent across the languages, as is illustrated in Figure 1. for the language-specific synsets linked to the ILI-record {drive}. Figure 1. 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.

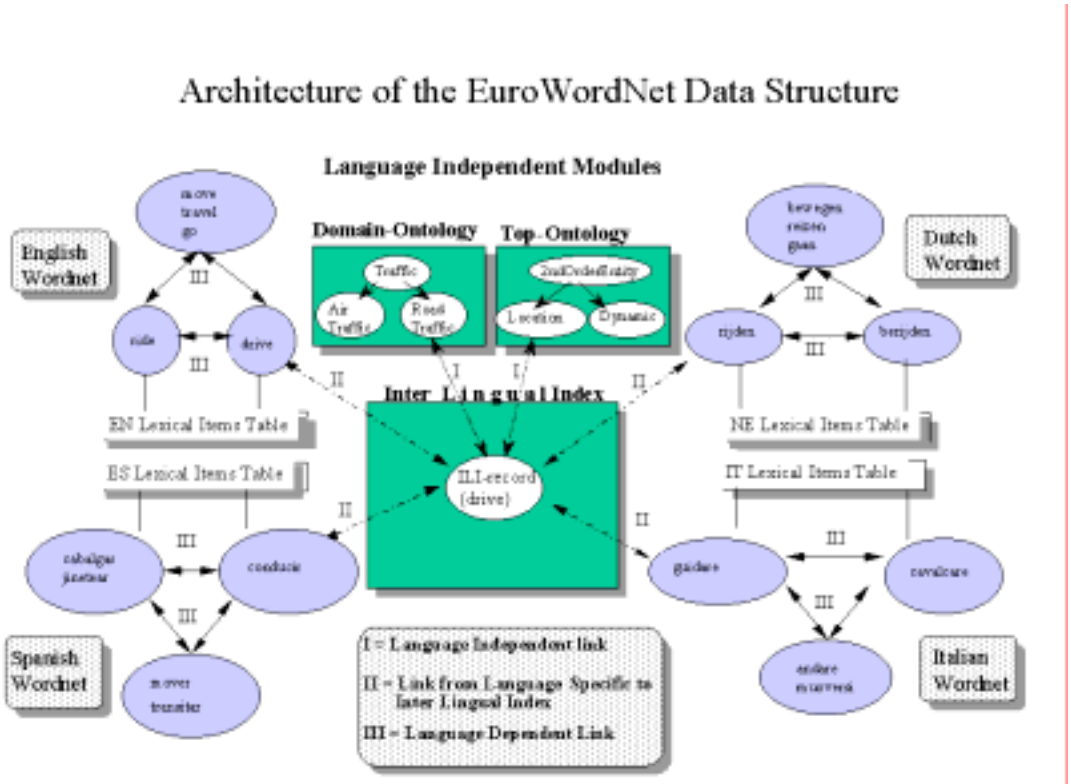


Figure 1. The global architecture of the EuroWordNet database.

The ILI is an unstructured list of meanings, where each ILI-record consists of a synset, a gloss specifying the meaning and a reference to its source. No relations are 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. The major advantage of this design is that both the language-specific relations and the equivalence relation do not have to be considered from a many-to-many perspective. A further discussion on the advantages and disadvantages of different multilingual designs and the ways of comparing the wordnets is given in [Vossen et al 1997a], and [Peters et al. fc.].

The separate construction of wordnets at different sites and the modular design of the EuroWordNet database provides a lot of flexibility but it also creates a major disadvantage. Because each site has a different starting point in terms of the quality and quantity of available lexical resources, tools and databases, the interpretation and coverage of the different wordnets may easily drift apart. Obviously, comparison and transfer as described above is only an option when the information in each wordnet is coded in a more-or-less compatible way. There are two specific compatibility issues at stake:

- the coverage of the vocabulary: the wordnets should globally deal with the same conceptual areas or domains
- the interpretation of the relations should be the same for all the sites.

To achieve maximal consensus, the wordnets are therefore developed top-down starting with a shared set of so-called Base Concepts which are classified using a common shared semantic framework. These Base Concepts are the most important meanings prevailing in the local wordnets and making up the core of the multilingual database. To get to grips with these meanings, they have been classified by a Top Ontology, specifically designed for that purpose. The Top Ontology is linked to the ILI as illustrated in Figure 1, thus providing some language-independent structuring of the ILI, together with the Domain Ontology:

- The Top Concept ontology is a hierarchy of language-independent concepts, reflecting important semantic distinctions, e.g. Object and Substance, Dynamic and Static;
- The Domain Hierarchy contains knowledge structures grouping meanings in terms of topics or

scripts, e.g. Traffic, Road-Traffic, Air-Traffic, Sports, Hospital, Restaurant;

In EuroWordNet, the Domain Ontology will only be implemented for illustration. The work has focused on the Top Ontology. Both the Top Concepts and the domain labels can be transferred via the equivalence relations of the ILI-records to the language-specific meanings. In Figure 1, the Top Concept *Motion* is for example directly linked to the ILI-record *drive* and it therefore indirectly also applies to all language-specific concepts related to this ILI-record. Via the language-internal relations the Top Concept can be further inherited to all other related language-specific concepts. In this way the wordnets are built using a common framework, where the lexicalizations around these Base Concepts may differ from language to language.

This document describes the selection of the Base Concepts and the Top Ontology that has been used to classify them. In the next section we will first further motivate the general top-down approach we have adopted. In section 3 we shortly describe the resources and methodologies of each site to clarify the different backgrounds and starting points. In section 4 we describe the technical procedure by which the set of common Base Concepts (BCs) has been established. Section 5 then describes a Top-Ontology of basic semantic distinctions, which has been developed to get to grips with these BCs. All BCs have been clustered in terms of these Top-Ontology Concepts providing a shared descriptive framework for the covered vocabulary. In Appendix II, the classifications of the Base Concepts is given. Electronic versions can be down-loaded from http://www.let.uva.nl/~ewn.

2. The general approach for building the wordnets

When defining the vocabulary we are faced with several conflicting requirements:

1. The vocabulary has to be generic: include all general word meanings on which more specific concepts depend and those meanings that are used most frequently.
2. The conceptual coverage across the different wordnets has to be the same: that is they should roughly contain the same areas of concepts.
3. The vocabularies should nevertheless reflect or at least respect language-specific lexicalization patterns.
4. There should be maximum freedom and flexibility for building the wordnets at the different sites: due to the different nature of the resources and tools there may not be one unified approach to build the wordnets which is best for all sites.

To achieve 2 we could simply take a particular set of synsets from WordNet1.5 as a starting point and make sure that these concepts are translated into the other languages and that language-internal relations are provided in these languages. However, this would endanger requirement 1 and 3, and perhaps also 4, for several reasons. First of all, the selection will be strongly biased by English and by the specific features of WordNet1.5 (including imbalances in the vocabulary of WordNet1.5). These are not automatically applicable to the other languages. What is more important is that we may miss typical lexicalizations and important meanings which are relevant to the other languages but which do not follow from the structure of WordNet1.5.

The assessment of the above requirements implies control at two levels: within each individual language and cross-linguistically. For these reasons we adopted a more complicated approach which will however establish a better common ground applicable to all the wordnets:

1. Each group separately defines a fragment of the vocabulary in the different local resources using the same criteria.
2. The local selections are then translated to equivalent WordNet1.5 synsets.
3. The sets of translations are compared to see how much overlap there is across the sites.
4. From this comparison a common set will be determined.
5. Each site adapts their selection to include the common set

After such a cycle the vocabulary will then be extended and the steps 1 through 5 are repeated.

What should then be the criteria for making these local definitions? The major conceptual criterion given above is that it should include all the word meanings that play a major role in the different wordnets and those meanings that are used most frequently. The latter is difficult to verify because

there are still no data on frequency of meanings. Instead of word meaning frequency, the selections can only be verified for word frequency. Fortunately, the former criterion can be satisfied by taking those meanings that exhibit most relations with other meanings and/or that occupy high positions in the hierarchies. There are several reasons for focusing on this group:

- These word meanings and their direct semantic neighbourhood form the core of the wordnets, on which the meanings of the other words depend.
- It is easier to extend the wordnets with more specific concepts when all the basic building blocks are present and well-defined.

Furthermore, early experiments in building some wordnet fragments showed that many problems in encoding relations are concentrated in a relatively small set of complex word meanings that strongly correlates with this set. Typically, words at the higher, more abstract levels of hierarchies, such as *object, place* or *change*, tend to be polysemous, have vaguely-distinguishable meanings and cannot easily be linked to other more general meanings. Furthermore, the available resources are often not very helpful either for these words (see [Vossen et al., fc.] for an extensive discussion of these problems). On the other hand, at the more specific levels (e.g. *tennis shoes*) meanings can be easily linked to a more general concept (*shoes*), also making the resources from which this information can be extracted more reliable.

To summarize, we see that the most important areas to create a generic semantic lexicon are also the most complex areas where resources are of little help. We therefore divided the building of the wordnets into two major phases:

- (mostly) manual construction of core-wordnets for a set of common Base Concepts and its direct semantic context.
- top-down extension of these core-wordnets, using (semi-)automatic techniques and relying on the information from the given resources.

In this way we can more effectively focus our manual effort on the more difficult and more important cases (also exchanging problems and solutions to achieve a maximum of consensus) and apply the automatic techniques to the areas of the resources which are more reliable. By starting off with a common set of Base Concepts we furthermore ensure that the cores of the wordnets are richly encoded and at least comparable: having the same conceptual coverage. On the other hand, there is sufficient freedom to fill in language specific lexicalizations and extensions in addition to this core. The rest of this deliverable is devoted to a further definition and characterization of the Base Concept Vocabulary. First, we will shortly specify the different backgrounds and, next, the local selections of Base Concept by each site and the common set of Base Concepts derived from these. This set is then further characterized using a Top-Ontology of basic semantic distinctions. The second phase, the Extension of the Base Concepts will not be discussed in detail in this paper. We will just acknowledge that the extensions will be based on the following general criteria:

- In general, the Common Base Concepts will be extended with their hyponyms. In a first phase of extension the most relevant hyponyms will be included. Relevance is based on their potential of further extendibility —i.e., hyponyms that have hyponyms themselves. The output will then be re-considered in order to add hyponyms of those Base Concepts which appear to be badly represented using such criterion —i.e., BCs which give raise to shallow hierarchies.
- A degree of common coverage of vocabulary across languages has to be ensured. We therefore compare and adapt the local coverage to achieve a reasonable level of overlap across the wordnets. This will be carried out in a similar way as is described for the Base Concepts in section 4.2 below.
- To further ensure that the wordnets include the generic parts of the vocabulary we will compare the selection of entries in EWN with the vocabulary covered in another EC-project Parole. In the Parole project morpho-syntactic lexicons for the most frequent words are built, where the frequency is derived from comparable corpora in several European languages.
- Finally, some specific vocabulary will be added for the domain of computers and software. The domain-terminology is necessary to test the EWN database for Information Retrieval tasks.

3. Local Resources and Methods

All the partners involved in EWN have a variety of lexical resources and tools at their disposal and had developed methodologies for performing their specific tasks. In [Vossen (ed.) 1997] a detailed account of such resources is presented. In the following paragraphs a summary of the main lexical resources is presented.

The University of Amsterdam (henceforth AMS) uses an object-oriented lexical database system [Boersma 1996] developed for the Sift-project (LRE 62030). The object-oriented treatment of the data makes it possible to efficiently manipulate lexicons, collections of entries, collections of senses or single entries and/or senses. Within the AMS LDB the following resources have been loaded for EuroWordNet:

- Celex Dutch lemma lexicon with basic syntactic information and corpus frequency information.
- WordNet 1.5
- The content of a lexical database VLIS (Van Dale Lexical Information System) provided by Van Dale.
- The Van Dale Dutch-English dictionary [Martin and Tops 1986].

The data from the Van Dale Lexical Information System (VLIS) has been used as input for developing the Dutch wordnet. The database contains the merge of several contemporary Dutch dictionaries published by Van Dale in recent years. The coverage of VLIS is as follows:

	nouns	*verbs*
Entries	63962	8822
senses	74678	14268

The Van Dale database is sense-oriented and contains, in addition to traditional information (such as definitions and usage codes), explicit semantic relations between word senses. Important semantic relations in VLIS are hyp(er)onymy, synonymy, antonymy, partitive and associative. The hyponymy-relations result in 1727 tops (1429 noun tops and 298 verb tops). As such it can be seen as a partially-structured semantic network similar to WordNet1.5.

At the Istituto di Linguistica Computazionale del CNR, in Pisa (PSA), three main sources for the Italian data are used:

- The Italian Monolingual Lexical Database (constructed from a number of sources).
- Italian/English Bilingual Lexical Database (constructed on the basis of the Collins-Giunti Italian-English dictionary)
- An Electronic Dictionary of Synonyms available at ILC-CNR.

Main figures for the Italian lexical database are the following:

	nouns	*verbs*
entries	24,635	5,546
senses	45,608	14,091

This database is enriched with a number of semantic relations between senses: hypernymy, meronymy, causation, verb_to_noun, adjective_to_noun. This monolingual LDB has been used as the main source of data for the Italian wordnet; the semantic relations, with the exception of the synonym and antonym relations, are extracted (when present) from this source. The size of the bilingual database is approximately 30,000 senses on each side (Italian-English, English-Italian).

The main sources used by Spanish group (FUE) are:

- Spanish Monolingual: DGILE- Diccionario General Ilustrado de la Lengua Española [Alvar Ed., 1987].
- English/Spanish and Spanish/English Bilinguals: VOX-HARRAP'S Esencial [VOX-HARRAP'S 1992] and VOX Advanced
- PIRAPIDES Verbal Database. Developed within the PIRAPIDES Project at the University of Barcelona.
- Several (partial) semantic (sense-based) taxonomies developed within the Acquilex project.

The figures for the monolingual dictionary are:

	nouns	*verbs*
entries	65,000	11,000
senses	105,000	24,000

The Pirapides database consists of 3600 English verb forms organized around Levin's Semantic Classes [Levin 1993] connected to WN1.5 senses. The database contains the theta-Grids specifications for each verb (its semantic structure in terms of cases or thematic roles), translation to Spanish forms and diathesis information.

The following resources are used by the University of Sheffield (SHE) for English:

	nouns	*verbs*	*adjectives*
Monolinguals			
LDOCE[1]	21400	7361	7333
COBUILD[2]	6566	6566	3490
Other Data			
CELEX	29494	8504	9185
COMLEX	21871	5660	8170

Given the available resources, each group developed different methodologies for selecting candidate nodes, extracting the relations (both internal and external) and linking each entry to the appropriate WN1.5 synsets. All methodologies combine automatic procedures with manual work.

In the case of AMS, the main source for both the entries and relations is the Vlis database. The relations that match the EWN relations have been copied to the EWN structure. The building of the Dutch wordnet then mainly consists of:

- verifying the copied relations and information, which can be confirmed, edited or removed.
- adding missing relations

For this manual process a special editor, so-called Surf-Editor, has been developed in the AMS LDB, that makes use of the fact that entries and senses are linked as hyper-text windows. Using this editor relations between multiple windows with activated senses can be edited, added or removed, while going from link to link (possibly in parallel for multiple resources). Only after the relations for the BCs have been coded, automatic techniques will be used to extract additional information from the definitions in monolingual dictionaries or translations in bilinguals. This information is compared with the information given or directly added when such information is missing.

At Pisa (PSA), it was decided to construct the Italian wordnet from a number of sources (at

[1] Procter et al. (ed.), 1987.
[2] Atkins S. (****)

least, at the upper level of the taxonomies) to overcome, to some extent, the idiosyncracies of a single dictionary and to provide a more objective perspective on the data. The starting point was the creation of the BCs using data from the 3 different sources mentioned above. However, an integration of different sources has also highlighted the differences between them and the inconsistencies found in dictionary data: e.g. word senses and synonyms vary from source to source. So a considerable manual effort was devoted to guarantee the quality of the selection.

For Spanish, an approach more closely related to WN1.5 was followed. The starting point was to take the two higher levels in WN1.5 hierarchy. First the WN1.5 synsets have been translated (using bilingual resources) and the basic semantic relations have been established (only hypernymy-hyponymy, synonymy, antonymy and causation in the first phase). This result has been used to extract the BCs for Spanish. In a second phase, additional taxonomies and monolingual resources are used to extract additional information and verify the results of the first phase.

Sheffield (SHE) takes a special position in the project because there is already a wordnet for English. The main task for SHE therefore consists of adapting WordNet1.5 by adding newly distinguished relations and improving the WordNet1.5 synsets that are used in the Inter-Lingual-Index for interlinking the wordnets (see [Peters et al., fc.] for details).

4. The Base Concepts

The main characteristic of BCs 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)

It should be noted that these criteria can not be applied in an absolute sense. To precisely measure the number of relations and the position in the hierarchy, these relations have to be established and finalized in the first place. All sites however use partially structured data that will be changed considerably during the project. The selections below should therefore be seen as global approximations of the set of BCs. Only in the case of the selection for English it was possible to use more sophisticated measurements because WordNet1.5 was available as a stable resource. To establish a minimal level of cohesion in approach and results for the individual selections of BCs, each group used these criteria as the main basis in one form or another, where the exact working out may differ due to the different starting points (see below). Additionally, some other criteria have been applied by some sites such as selecting all the members of the hypernym chain of any already selected BC or general frequency in sources (MRDs, corpora).

4.1. Definition of the local BCs

Following the above criteria, an initial set of noun and verb senses, grouped in synsets, has been selected for each language given the available resources.

For AMS, the Vlis hierarchy was sufficiently structured to extract information on the importance of concepts. First, the meanings with most relations have been selected, summing up to 15% of the total amount of relations in the database. For nouns this comes down to all meanings with more than 15 relations, for verbs to all meanings with more than 12 relations. The resulting set was further limited by restricting it to meanings occurring at a hierarchical depth of 3. This initial set was extended with:

- excluded tops which nevertheless have a lot of descendants; where tops are defined as meanings without a hyperonym relation;
- excluded hyperonyms of the words selected so far; to make sure that all meanings needed to anchor the most frequent ones are present in the set;

The Dutch set of BCs has been manually translated into WordNet1.5 equivalences. In 6 cases there was no good equivalent in WordNet1.5 for a Dutch BCs. In that case the Dutch BC was represented by the closest synset in WordNet1.5. In quite a few cases a single Dutch BC matched with several WordNet1.5 synsets. In that case, all the matching synsets have been generated. The reversed situation also occurred, although less frequently. In that case multiple Dutch BCs have been represented by a single WordNet1.5 synset.

In order to identify local Base Concepts, Pisa has used a semi-automatic procedure. A first list of lexical items was extracted automatically from the Italian monolingual LDB using as main criteria 1) the position (medium/high) in the taxonomy and 2) the number of relations with other lexical items (generally hyponyms). This set was then processed manually to meet the following objectives:

1. Overcome inconsistencies and lack of homogeneity of the data caused by nature of the sources and the automatic extraction techniques. For instance, if within the area of kinship terms the original extraction included 'husband' but not 'wife', the latter term was manually added.

2. Organize the data in terms of synonymy (i.e., grouping senses in synsets) and taxonomy.

The grouping of terms in synonyms was carried out semi-automatically. First, for each concept, information about potential synonymy was extracted automatically from the sources; then the resulting data were carefully evaluated and structured in synsets. The next step, namely the hierarchical organization of synsets, was performed manually after having realized that the application of automatic techniques to the existing sources was not useful to perform the task. This was due to several reasons, among them the following: (a) Many terms are defined in the dictionary by means of synonyms; (b) Many terms are defined by means of potential hypernyms carrying a low semantic value —e.g. "atto", "effetto", "modo" (act, effect, manner). The list of BC synsets identified was then mapped to WN1.5 in order to establish cross-language lexical equivalences. Some problems were identified at this stage: e.g., it was difficult to identify the correct WN sense or to distinguish between close senses - often one Italian word-sense was mapped to more than one WN1.5 sense, sometimes also vice versa.

For Spanish, two complementary main sources were used in the case of nouns: 1) An extended taxonomy of Spanish obtained from the monolingual Vox dictionary and 2) A manual translation of the two higher levels of WN1.5. For verbs the main source was 3) the Pirapides database, already connected to WN1.5. Two additional sources were used as well: 4) Frequency counts of words in the definition (and examples) field of the monolingual DGILE and 5) Frequency counts of words in LEXESP (a 3 Mw balanced corpus of Spanish). The main criteria used in this case were:

1. A selected word is a translation of either a top concept or a direct hyponym of a top concept in WN1.5, and either (2) or (3):

    2. It occurs as genus word in the DGILE monolingual MRD 5 or more times.

    3.It shows a high frequency of occurrences in corpora: either (3.1) or (3.2):

        3.1. It occurs 50 times or more in the DGILE definition corpus.

        3.2. It occurs 100 times or more in LEXESP

SHE, for English, has used the notion of conceptual density as the main criterion, for which three measures have been considered:

a) a node's total number of hyponymic descendants;

b) a node's mean branching factor (mean number of hyponyms in WN1.5);

c) a function of a) and the node's relative position in the hyponymic hierarchy.

After empirical investigation definition c) proved to be the most promising, and the result was computed in the following way:

$$\frac{\text{total no. of descendants}}{\text{level of concept / total no of levels of the chain including the concept}}$$

Extracting the 20% topmost values for nouns yielded 1090 distinct noun synsets. For verbs the algorithm resulted in 197 distinct verb synsets.

4.2. Deriving a common set of Base Concepts

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 coincided. In so far they did not, we had to apply special measures to achieve a reasonable common set, which is a condition to make the cores of the wordnets compatible.

Only 30 BCs are part of all selections (24 noun synsets, 6 verb synsets). This is extremely low considering the uniformity of the criteria. There can be several explanations for this:

1. The individual selections are not representative enough.
2. There are major differences in the way meanings are classified, which have an effect on the frequency of the relations.
3. The translations of the selection to WordNet1.5 synsets are not reliable
4. The resources cover very different vocabularies

The second explanation is acceptable and is inherent to our approach where each wordnet represents an autonomous language-internal network. Differences in the way meanings are classified will shows up when the wordnets are compared. This may lead to a restructuring and to a more coherent set of important, classifying Base Concept in the local wordnets. The fourth explanation is not likely to apply to general words and meanings. Since all sites use contemporary monolingual resources we do not expect that the core vocabularies differ a lot in coverage.

With respect to the first and third explanations we took some specific measures. First of all, the individual sets may be too small to be representative but the merge of the sets may be sufficiently comprehensive. Instead of the total intersection of concepts we therefore took all synsets selected by two sites.

Table 1a: Intersections for Nouns in terms of WordNet1.5 synsets

	AMS	*FUE*	*PSA*	*SHE*
AMS	1027	103	182	333
FUE	103	523	45	284
PSA	182	45	334	167
SHE	333	284	167	1296

Table 1b: Intersections for Verbs in terms of WordNet1.5 synsets

	AMS	*FUE*	*PSA*	*SHE*
AMS	323	36	42	86
FUE	36	128	18	43
PSA	42	18	104	39
SHE	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. Given this set of common Base Concepts the local selections can be divided into:

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

In addition, a third subset of BCs is assigned to each site:

- missing, i.e. this synset has been selected by at least two other sites, so it must be added as BC in the local selection

The result for each group is given in the next table

Table 2: Selected and Rejected Base Concepts Nouns

Nouns	*Proposed*	*Selected*	*Rejected*	*Missing*
AMS	1027	429	598	265

FUE	523	323	200	371
PSA	334	239	95	455
SHE	1296	594	702	100
Union	2287	694	1595	
Verbs	***Proposed***	***Selected***	***Rejected***	***Missing***
AMS	323	126	197	51
FUE	128	72	56	105
PSA	104	63	41	114
SHE	236	132	104	45
Union	573	177	398	

These tables illustrate the fact that in the case of AMS nouns, for instance, from 1027 candidates (local BCs) 429 were selected (as being members of at least another selection) and 598 were rejected. The last column says that 265 senses, belonging to the common BCs were missing in the local selection and thus have to be added to the AMS selection. The selection of the common BCs (CBCs) thus resulted in a set of missing nouns and verbs for every language. Each group tried to represent the missing BCs as far as possible. It may be the case that there are no exact equivalents in the local language for a common BC. In that case we tried to include the closest concept in the local selection.

4.3. Tuning of the initial Common Base Concepts

This first BC selection, described in the two previous sections, was based on the local wordnets by taking into account their intersections without any examination or evaluation of the selected synsets. Because of the extremely low intersection of the local selections we decided to add a second selection phase in order to 'tune' the original BC set and extend it into a conceptually more homogenous group of concepts. This extension has been based on the synsets, which originally had been rejected in the first phase.

One explanation for the low intersection given above was the unreliability of the translations. As described in [Vossen et al., fc] and [Peters et al., fc] the degree of polysemy of WordNet1.5 is much higher than in traditional resources. For example, a verb such as *clean* has 19 different senses in WordNet1.5, whereas traditional dictionaries only give one general sense. A danger of this extreme sense-differentiation is that a single sense in the traditional resources may match with several synsets in WordNet1.5. As a result of this, it is not unlikely that, in many cases, different WordNet1.5 synsets have been chosen for language specific synsets, which are in fact equivalent.

To measure the possible impact of this type of mismatching, we checked to what extent the rejected and selected BCs represent different senses of the same entries. The following table gives an overview of the matches between Rejected Concepts (RCs) and BCs at the word level:

Table 3: Multiple rejected senses of the same word.

Nouns	***Entries***	***Synsets***
RCs sharing one or more word forms with BCs	303	529
RCs not sharing any word form with BCs	87	194
Total	390	723

Verbs	***Entries***	***Synsets***
RCs sharing one or more word forms with BCs	158	285
RCs not sharing any word form with BCs	50	124
Total	208	409

From these we selected all RCs that represent different senses of the same entries (either RCs or RCs and BCs). This set has been further limited by the following constraints:

1. Only words shared by at least four synsets, RC or BC, have been included in the evaluation. In other words: at least four senses of a word must be involved. A check has been made to ensure that the rejected synsets belonging to these sense groups did not all originate from only one language specific wordnet, but have a more or less even distribution over the different language sites.
2. We have focussed on synsets that have more than average relations (19.49 for the BCs in WordNet1.5). This includes the relations for RCs as separate synsets, but also the merged relations of RCs that are very close.

Next we have carried out a manual check of all these cases to see whether the RCs have been rejected because of a mismatching of translations from language specific concepts to WordNet1.5 senses. RCs have been reselected if:

a) Their meaning is more central or basic than a selected BC.
b) They have more than average number of relations.
c) They can be merged with another selected or rejected BC because they are very close in meaning.
d) They exhibit a regular polysemy relation with a selected or rejected BC: e.g. metonymy, diathesis alternations. For a further discussion on the identification of sense relations in terms of generalization and metonymy see [Peters et al., fc.].

To measure the closeness of senses of entries, a metric was applied to nominal RCs which had been developed by [Agirre and Rigau 1996] for computing conceptual distance between RC and BC WordNet1.5 nodes. This measure takes into account the length of the shortest paths that connects the concepts involved, the depth of the hierarchy, and the density of the distribution of concepts in the hierarchy. If an RC-BC pair was found to be conceptually very close the RC synset was selected. A conceptual distance threshold value of 0.3 was considered to be the best criterion for selection.

An RC has not been reselected if:

a) In order to maximize the coverage of the BC set direct RC hyponyms of existing BCs have principally not been selected unless they were judged strong enough candidates for inclusion.
b) Noun basic-level concepts [Rosch 1996, Lakoff 1987] such as {bed}, {wheel}, {shoe}, {window}, {glass}, {eye} and {soup} represent a level of lexicalisation which is considered too specific for our selection purposes, and have not been selected.
c) Nominal taxonomic terms within the field of biology have not been selected as new BCs. They have very specific technical meanings, and are subsumed by the BC `group'.

The selection of RCs involving the methods described above resulted in an extension of the BC set with 133 noun and 62 verb synsets.

Adding these senses, resulted in the final set of common BCs of 1059 synsets, representing 796 nominal BCs and 263 verbal BCs. Each group tried to represent the missing BCs (either in the first selection or added in the fine-tuning) 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 4: Number of Common BCs represented in the local wordnets

1059	*Local Synsets Related to CBCs*	*Eq_synonym Relations*	*Eq_near_ Synonym relations*	*CBCs Without Direct Equivalent*
AMS	992	725	269	97
FUE	1012	1009	0	15
PSA	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, 13 of which not in two wordnets:

Table 5: BC4 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**
material#3; matter#5	Psychological feature#1
spatial property#1; spatiality#1	

Only 2 CBCs could not be represented in all 3 wordnets: *body covering* and *plant part*. Apparently, they have only been selected because they are important levels in wordnet or closely related to other senses which are selected. 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. The following complex equivalence relations have been created to the CBCs:

Table 6: Local senses with complex equivalence relations to CBCs

	AMS	*FUE*	*PSA*
Eq_has_hyperonym	61	40	4
eq_has_hyponym	34	14	20
Eq_has_holonym	2	0	
Eq_has_meronym	3	2	
Eq_involved	3		
Eq_is_caused_by	3		
Eq_is_state_of	1		

Here are some examples of complex relations in the Dutch wordnet:

{ongelukkig#1}, Adjective (unhappy)

Eq_is_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

{wonen#1}, Verb, (to live)

Eq_involved home#1, 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 illustarted 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 1059 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 pattern that is 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.

5. Top-Ontology clustering of the Base Concepts

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. There is no common, a priori agreement how to build an ontology. In fact there is no agreement on what an ontology is (collections of related objects so different as CYC [Lenat and Guha, 1990], Generalized Upper Model [Bateman et al. 1994] or WordNet1.5 [Miller et al. 1990] are considered ontologies). [Gruber 1992] therefore uses a pragmatic definition of an ontology: "an explicit specification of a conceptualization", i.e. a description of the concepts and relationships that can exist for an agent or community of agents. He points out that what is important is what an ontology is used for. The purpose of an ontology is enabling knowledge sharing and reuse. In that context, an ontology is a specification used for making ontological commitments. This definition can, of course, include the frequent forms of a taxonomic hierarchy of classes or a thesaurus, but also structures including and using more sophisticated inference mechanisms and in-depth knowledge about the world (or about the involved domain).

Ontologies differ in their scope (general or domain specific), in the granularity of the units (just terminological labels or units owning more or less complex internal structure), in the kind of relations that can be defined between units, and in the more or less precise and well defined semantics of the units and relations (inheritance and other inference mechanisms). Further on, Gruber distinguishes between Representation ontologies and Content ontologies. The former provide a framework but do not offer guidance about how to represent the world (or the domain) while the latter make claims about how the world should be described.

[Gangemi et al. 19??] discusses several approaches for building ontologies based on most of these distinctions. They pay attention, basically, to the order of selection of the candidate nodes: a top-down approach, starting from domain-independent top-nodes, that seems to be more adequate for general ontologies, a bottom-up approach that tries to induce more general behaviour from local (mostly terminological) nodes or an hybrid approach (the ONIONS methodology in their case) that tries to take profit of both previous ones.

However, not only the direction of selection is important for deciding the building strategy. Different approaches can be followed for filling information. A stepwise refinement approach, based on a cascade of enrichment processes: first selecting the candidate nodes to form a simple list of names, then establishing in successive cycles relations between them, and, finally, filling the information

owned by each node. Of course, some of the refining cycles can be performed in not predefined order and sometimes in parallel. An alternative approach consists of starting with an initial node (or a small set of initial nodes), filling this node with all available information about it, establish all the relations involving this node and proceed recursively with each of the nodes related to it. The approach to be selected depends largely on the characteristics of the ontology to be built, i.e. domain, size, content, granularity, intended use, and so on.

5.1. Starting points for the EuroWordNet Top-Ontology

As explained in the introduction, the EuroWordNet database consists of separate language-specific modules (as autonomous systems of language-internal relations), which are linked through an Inter-Lingual-Index. The Inter-Lingual-Index (ILI) is an unstructured fund of synsets (mainly taken from WorNet1.5), the so-called ILI-records. Language-specific synsets linked to the same ILI-record are assumed to be equivalent. The ILI-records further give access to all language-independent knowledge, among which a Top Ontology of fundamental semantic distinctions. This language-independent information can be transferred via the ILI-record, which is assigned to it, to all the language specific synsets that are linked to it. This has been explained in the introduction in Figure 1, where local meanings linked to the ILI-record {drive} will indirectly be classified by the TCs linked to this ILI-record. 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 will enforce 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 verify 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, as is illustrated in Figure 1 above. 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. The most important purpose of the top-ontology is to provide a common starting point and high degree of compatibility across the wordnets for the BCs. 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.[3] It is therefore important that the top-ontology incorporates semantic distinctions that play a role in linguistic approaches rather

[3] 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.

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].[4]

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 4 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 are more like semantic features than like 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 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*.

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

5.2. The EuroWordNet Top-Ontology

The current ontology (version 1) 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.

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. continue, occur, apply

3rdOrderEntity
: An 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, though, 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 is 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 event can occur in several different places, not only at different times but also at the same time. 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. As described in [Alonge et al., fc.] this distinction cuts across the different parts of speech in that:

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

With respect to the BCs we therefore also see that all three parts-of-speech can be classified below the 2ndOrderEntity node. The actual distribution over the different parts of speech is shown in the next table:

Table 7: Total Set of shared Base Concepts

	Nouns	*Verbs*	*Total*
1stOrderEntities	491		491
2ndOrderEntities	272	263	535
3rdOrderEntities	33		33
Total	796	228	1059

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[123]
  - Artifact[144]

Form[0]
- Substance[32]
  - Solid[63]
  - Liquid[13]
  - Gas[1]
- Object1[62]

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]

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.

5.2.1. 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).

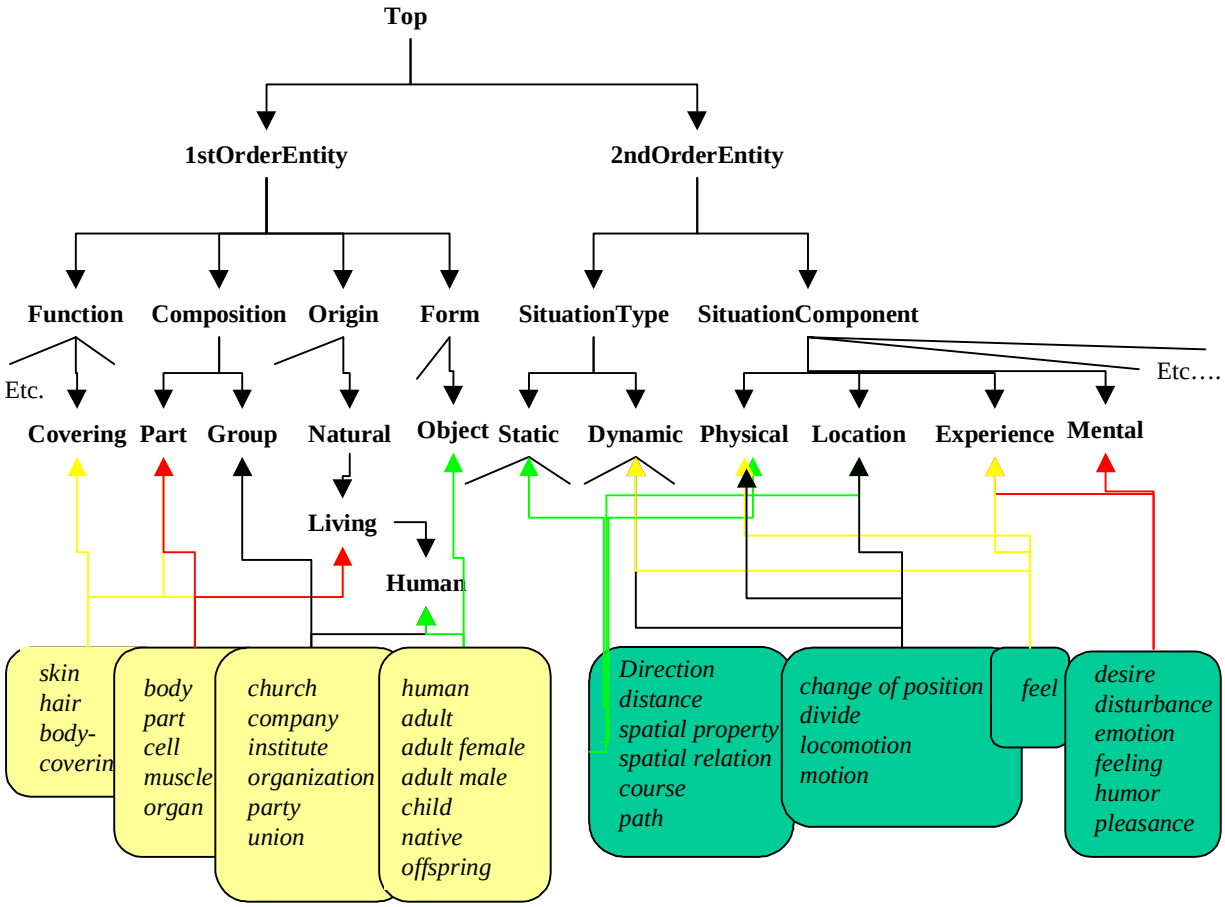
The main-classes are then further subdivided, where the subdivisions for Form and Composition are obvious given the above definition, except that Substance itself is further subdivided into Solid, Liquid and Gas. In the case of Function the subdivisions are based only on the frequency of BCs having such 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:



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.[5] 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.

[5] 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.

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.

5.2.2. 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.

5.2.2.1. 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: 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 to 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*).

5.2.2.2 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., fc.] for a further discussion on accessing the different modules in the database.

6. Conclusions

In this document we have described how we control the building of the separate wordnets at separate sites, where there has to be a maximum of flexibility, and still compatible results. On the one hand we want to allow for the development of unique language-specific wordnets, using different tools and methodologies, and on the other hand, we need to ensure that the same vocabulary is covered and the same decisions are made across the different sites. We therefore developed a top-down approach where the building is divided into two phases: 1) covering a shared set of common Base Concepts, 2) extending from these Base Concepts using semi-automatic techniques.

The Base Concepts represent the shared cores of the different wordnets, where we try to achieve a maximum of consensus and overlap. Still, the local wordnets can differ in the exact way in which the vocabulary is lexicalized around these Base Concepts. We further specified the definition and selection of the Base Concepts. The main criterion has been the relevance of meanings for the local wordnets. This relevance has been measured mainly in terms of the number of relations and the position in the hierarchy. The local selections have been translated to WordNet1.5 synsets and merged into a shared set of concepts. This set has been critically assessed and evaluated which resulted in a final set of 1059 Common Base Concepts.

To get to grips with the Base Concepts they have been classified using a Top Ontology. The Top Ontology provides a language-independent structuring of the Base Concepts in terms of 63 fundamental semantic distinctions. This classification is used as a common frame-work to further guide the encoding of the language-internal relations at each site.

References

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

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

[Bateman et al. 1994] Bateman, J., B. Magnini and J. Rinaldi (1994) *The Generalised Upper Model*. In Proceedings of ECAI 1994.

[Boersma 1996] Boersma, P. (1996) The Sift Lexical Database, Sift LRE 62030, Deliverable D10b, University of Amsterdam. Amsterdam

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

[Gangemi et al. 19??] Gangemi, A., G. Steve and F. Giacomelli. (19??) *Onions: An Ontological Methodology for Taxonomic Knowledge Integration.*

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

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

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

[Levin 1993] Levin, B. (1993) *English Verb Classes and Alternations*. University of Chicago Press. Chicago.

[Lyons 1977] Lyons, J. (1977) *Semantics.* Cambridge University Press. London.

[Martin and Tops 1986] Martin W. and J. Tops (1986) *Groot woordenboek Engels-Nederlands.* Van Dale Lexicografie. Utrecht.

[Miller et al. 1990] Miller, G., R. Beckwith, C. Fellbaum, D. Gross and K. Miller. (1990) *Five Papers on WordNet.* CSL Report 43. Cognitive Science Laboratory. Princeton University.

[Procter 1987] Procter, P et al. (Eds) (1987) *Longman Dictionary of Contemporary English.* Longman, Harlow and London.

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

[Pustejovsky 1995] Pustejovsky J. (1995) *The Generative Lexicon*. The MIT Press. Cambridge, MA.

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

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

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

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

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

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

[Vossen (ed.) 1997] Vossen, P. (Ed.) (1997) *Encoding the Semantic Relations for basic Nouns and Verbs*. Deliverable D010D011. LE-4002. Amsterdam.

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

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

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+Artifact
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+Communicat
ion
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+Communic
ation
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+Co
mmunication
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+Communicati
on
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+Communic
ation
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+C
ommunication
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+M
anner
blow 2: 00647048-n
BoundedEvent+Agentive+Physical+Location+Purpose+Po
ssession
get rid of 2: 01267839-v
BoundedEvent+Agentive+Physical+Location+Purpose+So
cial+Manner
stroke 3: 00329906-n
BoundedEvent+Agentive+Physical+Purpose+Communicat
ion
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+Soci
al
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+Soci
al+Manner
declaration 2: 04390828-n
BoundedEvent+Agentive+Purpose+Communication+Usag
e+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;Recreation
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

	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