Ontological Knowledge Enhancement in EcoLexicon

Juan Carlos Gil-Berrozpe, Pilar León-Araúz, Pamela Faber

University of Granada

Department of Translation and Interpreting, Buensuceso 11, 18071 Granada, Spain E-mail: jcgilberrozpe@ugr.es, pleon@ugr.es, pfaber@ugr.es

Abstract

Contemporary research has focused on how concepts are represented and organized in the mind, leading to neurocognitive theories such as grounded cognition or embodied cognition. These theories have greatly influenced further studies in linguistics and terminology. In this way, conceptualization, categorization, and knowledge organization are the foundation of cognitive-oriented terminology theories which highlight the relevance of situated knowledge structures, such as Frame-based Terminology. Accordingly, the practical application of Frame-based Terminology is EcoLexicon, a dynamic terminological knowledge base on environmental science. Concepts in this terminological resource are domain-specific within the Environmental Event, a model that interrelates concepts by assigning them different roles. However, the Environmental Event does not include specific category types to annotate these concepts ontologically. Therefore, this paper presents a process of ontological knowledge enhancement in EcoLexicon. This process was mainly based on the categorization of its concepts in semantic classes with a multidimensional approach. As a result, EcoLexicon was ontologically enhanced not only in terms of this categorization, but also through a redesign of the conceptual categories module, which involved modifying the existing category hierarchy and implementing new features focused on describing the combinatorial potential of concepts and categories (i.e. the conceptual combinations function and the ontological view).

1. Introduction

According to classic theories of cognition, knowledge representations are amodal data structures located in a semantic memory that is completely isolated from the modal systems of the brain for perception, action, and introspection (Barsalou, 2008). However, contemporary theories of cognition, including grounded cognition (Barsalou, 2010; Kiefer & Barsalou, 2013) and embodied cognition (Gallese & Lakoff, 2005; Martin, 2007; Meteyard et al., 2012), propose a more interrelated depiction of knowledge in our minds.

Grounded cognition considers that factors such as the environment, situations, the body and simulations are essential for contextualizing the cognitive representations in the brain's modal systems (Barsalou, 2010). Likewise, embodied cognition implies that the body is the main grounding mechanism and that all cognitive processes depend on perception and action (Meteyard et al., 2012). In line with this, concepts are embodied in the sense that their conceptual features are represented in sensory and motor brain areas based on experience (Kiefer & Pulvermüller, 2012). Not surprisingly, every discipline with a cognitive perspective pays attention to how concepts are represented and organized in the mind (Mahon & Caramazza, 2009) or, in other words, to how conceptual information is categorized.

These grounded or embodied approaches to conceptualization are particularly relevant to the fields of linguistics and terminology because of the cognitive shift (Faber, 2009) in these disciplines over the last decade. This cognitive shift has specifically affected the study of terminology in relation to specialized knowledge representation, category organization and conceptual description. Not surprisingly, terminology is a discipline that combines linguistic and cognitive facets, since terms are linguistic elements which carry conceptual meaning within the framework of specialized knowledge texts (Faber, 2009). As such, lexicographic and terminological resources should draw on various aspects or details coming from psychological studies.

Accordingly, cognitive-based theories of terminology are also inspired in contemporary theories of cognition. Thus, they claim that specialized concepts are not activated in isolation, but are typically contextualized in background situations and events (Faber & San Martín, 2010). For instance, when perceiving an entity, people also perceive the space where it is located, including the agents, patients or events affecting it. Moreover, brain-imaging experiments have confirmed that simulations of potential actions are greatly involved in the conceptualization of entities and events, even including those which are mentioned in specialized language texts (Faber et al., 2014).

Because of the influence of cognition in terminology, it is necessary to develop or enhance the ontological information displayed in terminological resources so as to offer more accurate representations of concepts and their descriptions. This would lead to a more expressive formal ontology, which would not only benefit human users by facilitating knowledge representation and acquisition, but also non-human users by offering a higher degree of interoperability and usefulness. In most cases, this process starts by structuring the knowledge contained in the resource in a given manner, and this is the point where categorization plays a key role. In fact, classifying knowledge through categorization is inevitable, because any concept can be included in a set of hierarchically-organized categories (Murphy & Lassaline, 1997), which can range from general to specific levels.

In this context, this paper addresses a process of ontological knowledge enhancement in EcoLexicon¹, a terminological knowledge base on environmental science. This process was mainly based on the categorization of its concepts in semantic classes with a multidimensional approach. As a result, EcoLexicon was ontologically enhanced not only in terms of this categorization, but also through the redesign of the previous conceptual categories module, which involved modifying the category hierarchy and implementing new features (i.e. the conceptual combinations function and the ontological view).

¹ http://ecolexicon.ugr.es/en/index.htm

2. Conceptual categorization of environmental knowledge

Neurological characteristics such as conceptualization, categorization, and knowledge are the foundation of Frame-based Terminology organization (FBT). \mathbf{a} cognitive-oriented terminology theory which highlights the relevance of situated knowledge structures represented as frames (Faber, 2015). FBT combines specialized knowledge representation with cognitive linguistics and semantics, taking aspects from both psychological and linguistic models. Frames are the cornerstone of FBT, and they are usually defined as the knowledge structures which contain information about the conceptual level and which relate entities and events associated with a particular scene or situation from human experience (Faber, 2015). Accordingly, any scientific or technical text contains specialized knowledge units that activate domain-specific semantic frames that are linked to the domain and to the user's background knowledge.

FBT has its main practical application in the form of a terminological resource: EcoLexicon (Faber et al., 2016). EcoLexicon is a dynamic terminological knowledge base on environmental science that provides a wide range of information about each of its entries, including conceptual, linguistic, phraseological, and multimodal aspects. EcoLexicon currently contains approximately 4,500 environmental concepts and 23,500 terms distributed in seven languages (English, Spanish, German, French, Dutch, Modern Greek, and Russian), with plans to include terms in Chinese, Portuguese, and Arabic. In addition, one of the most important functionalities in EcoLexicon is its general view (Figure 1), where conceptual networks are displayed and show how concepts are interrelated through different semantic relations (generic-specific, part-whole, and non-hierarchical relations).

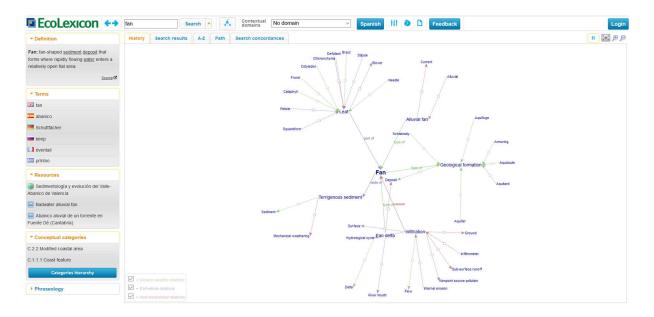


Figure 1: General view of EcoLexicon.

2.1 Environmental Event

According to FBT, conceptual networks are based on an underlying domain and on a closed inventory of both hierarchical and non-hierarchical semantic relations (Faber et al., 2009). These were the main premises used when building EcoLexicon, and the targets were conceptual relations and the combinatorial potential of concepts, extracted from corpus analysis.

In EcoLexicon, knowledge can be accessed from general to more specific relational structures. The most basic level is the Environmental Event (EE). In this frame, general categories of environmental entities are linked by predicates codifying the states, processes, and events in which the entities can take part (Faber, 2015). As stated by León-Araúz et al. (2012), the EE contains basic meanings that relate concepts, roles, and categories pertaining to general environmental knowledge. Moreover, the EE also links generic categories at the superordinate level and provides the basis for subframes that can be used to restrict contextual information to what is most relevant.

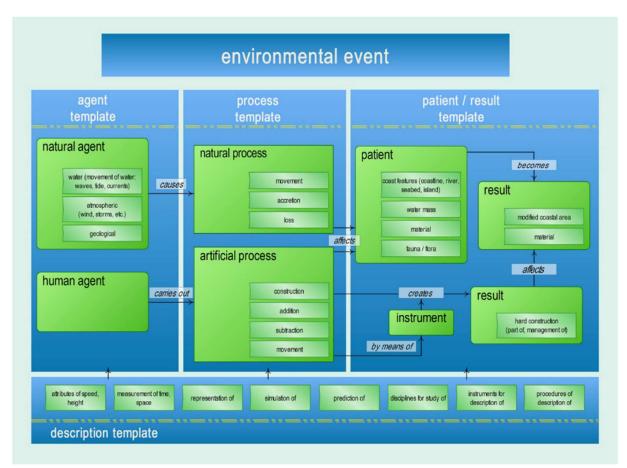


Figure 2: Environmental Event.

As shown in Figure 2, the Environmental Event has two types of AGENT that can initiate processes, i.e. NATURAL AGENTS (inanimate) and HUMAN AGENTS (animate).

On the one hand, natural forces (e.g. water movement) cause NATURAL PROCESSES (e.g. river erosion) in specific locations, commonly regarded as PATIENTS (e.g. riverbed) which, as a RESULT, may suffer alterations (e.g. deterioration, modification of size or shape). On the other hand, humans can also carry our ARTIFICIAL PROCESSES (e.g. construction) to alter the EFFECTS normally caused by natural processes (e.g. protection), or to create new effects through the use of certain INSTRUMENTS (e.g. defence structures).

Nevertheless, the conceptual representation of environmental knowledge cannot be achieved simply by assigning these generic semantic roles to concepts as if all of them would belong to a universal type of event (León-Araúz et al., 2012). In fact, contextualization has to be taken into account, because the way in which a concept interacts with other concepts can influence its categorization (Evans & Green, 2006). For this reason, the EE was originally used as a macrostructure for the further design of context-dependent microstructures (e.g. coastal engineering, meteorology, oceanography).

In recent years, the content of EcoLexicon has widely expanded, including a large quantity of conceptual and semantic information that has allowed us to interrelate all of its content, and thus go beyond the specific cases observed in the original EE. Because of this expansion in conceptual meaning, the need for an enhanced ontology of environmental categories has become apparent, since the EE does not include specific category types to annotate all environmental concepts ontologically, but only semantic roles. For this reason, we carried out an in-depth categorization process of all concepts in the database, a revision of the ontology underlying EcoLexicon, and the implementation of new features to its conceptual categories module, which will be explained in the following sections.

2.2 Conceptual categorization process

An ontology is usually regarded as a database describing the concepts of a knowledge field, their properties or characteristics, and how concepts are related to each other (Weigand, 1997). Moreover, ontologies are often organized as classification hierarchies and tend to be as universal as possible so that they can be used and reused for different applications. Such hierarchies tend to position the three most basic ontological categories at the top level: ENTITIES or OBJECTS, PROCESSES or EVENTS, and ATTRIBUTES or PROPERTIES (Mahesh & Nirenburg, 1995; Moreno-Ortiz & Pérez-Hernández, 2000).

In this context, various ontology-based projects for categorizing environmental knowledge have already been carried out, such as the Environmental $Ontology^2$

² http://www.obofoundry.org/ontology/envo.html

(ENVO) (Buttigieg et al., 2013, 2016). More specifically, ENVO defines itself as "a community-led, open project which seeks to provide an ontology for specifying a wide range of environments relevant to multiple life science disciplines and, through an open participation model, to accommodate the terminological requirements of all those needing to annotate data using ontology classes" (Buttigieg et al., 2013). Although this project was initially focused on the representation of biomes, environmental features, and environmental materials, it has been continuously expanding to include ontological information related to a multitude of interrelated fields (Buttigieg et al., 2016).

In a similar way, the conceptual categorization process in EcoLexicon followed the premises behind ENVO's ontological reasoning by adapting the conceptual categories and hierarchies to the specific needs of the environmental knowledge contained in EcoLexicon. Because of the dynamism of environmental sciences (León-Araúz et al., 2012), it was essential to take into account the multifaceted nature of concepts, as they can belong to more than one category depending on their salient features (Kageura, 1997). For this reason, the conceptual categorization process was carried out from a multidimensional perspective.

A series of semantic classes belonging to different top-down categorization levels was established to determine degrees of specificity (Murphy & Lassaline, 1997) and conceptual similarity (Hahn & Chater, 1997), so that every concept could be tagged with a category showing its interrelation with ontologically-similar elements. These semantic classes were mainly based on concept definitions and on the contextual information in the EcoLexicon corpus, but they were also contrasted with the ontological classes found in ENVO (Buttigieg et al., 2013, 2016). Consequently, an enhanced category system for EcoLexicon was established and hierarchically organized (Figure 3).

-			
Process	Loss		
Process	Method		
Process	Movement		
Process	Movement	Earth / Soil movement	
Process	Movement	Energy movement	
Process	Movement	Fluid movement	
Process	Movement	Fluid movement	Water movement
Process	Movement	Transport	
Process	Movement	Wave	
Process	Movement	Wind movement	
Process	Phase		
Process	Phase	Phase of cycle	
Process	Phase	Phase of treatment	
Process	Phenomenon		
Process	Phenomenon	Atmospheric phenomenon	
Process	Phenomenon	Atmospheric phenomenon	Precipitation
Process	Phenomenon	Optical phenomenon	

Figure 3: Example of the category hierarchy.

In this way, the 4,500 concepts in EcoLexicon were classified in 152 categories, distributed in five categorization levels. To begin with, the most general level is composed of the three starter ontological categories (Mahesh & Nirenburg, 1995; Moreno-Ortiz & Pérez-Hernández, 2000):

A: ATTRIBUTE – properties of entities and processes

E: ENTITY – physical and mental objects

P: PROCESS – events extending over time and involving different participants

However, depending on the ontological nature of concepts, they can be subclassified in up to five levels of specificity, as can be seen in the category hierarchy involving CREATION concepts:

E: ENTITY

E-1: CREATION
E-1.1: ARTIFACT (e.g. dc bus)
E-1.1.1: CONDUIT (e.g. duct)
E-1.1.2: CONTAINER (e.g. sedimentation tank)
E-1.1.3: INSTRUMENT (e.g. centrifugal pump)
E-1.1.3.1: MEASURING INSTRUMENT (e.g. accelerometer)
E-1.1.3.2: RECORDING INSTRUMENT (e.g. albedograph)
E-1.1.3.3: SAMPLING INSTRUMENT (e.g. automatic sampler)
E-1.1.3.4: TRANSFORMING INSTRUMENT (e.g. solar cell)
E-1.2: SOFTWARE (e.g. computer application)
E-1.3.1: BUILDING (e.g. oil refinery)
E-1.3.2: DEFENSE STRUCTURE (e.g. reef breakwater)

Additionally, those concepts with a multidimensional nature (Kageura, 1997) were classified in as many categorization hierarchies as necessary, depending on the salient features observed in their definitions and in the corpus. For instance, one of the most multifaceted concepts is *port*, which was classified according to four categories:

- Concept: port
- **Definition (from EcoLexicon):** place along a river or seacoast that gives ships and boats protection from storms and rough water, and where ships can load and unload cargo. It can be natural or artificial.
- Conceptual category:
 - E-1.3: STRUCTURE
 - E-4.1: ARTIFICIAL GEOGRAPHIC FEATURE
 - $\circ\quad E-4.2: \text{ NATURAL GEOGRAPHIC FEATURE}$
 - E-12.1.2: FACILITY

Figure 4 shows a fragment of the categorization table that was used to summarize the classification process. The first column contains the concept analyzed; the second column indicates whether the concept is multidimensional; the third column describes the number of categories applied to a single concept; and the remaining columns contain the top-down categories applied to each concept.

sheet pile	NO	FIRST CATEGORY:	Entity	Part	Part of structure
taiga	NO	FIRST CATEGORY:	Entity	Geographic feature	Natural geographic feature
stem	NO	FIRST CATEGORY:	Entity	Part	Part of lifeform
thallus	YES	FIRST CATEGORY:	Entity	Part	Part of lifeform
utatius	TES	SECOND CATEGORY:	Entity	Part	Part of lifeform
slope	NO	FIRST CATEGORY:	Entity	Part	Part of landform
continental slope	continental slope YES	FIRST CATEGORY:	Entity	Space	Layer
continental slope	125	SECOND CATEGORY:	Entity	Part	Part of landform
grain size	NO	FIRST CATEGORY:	Attribute	Physical attribute	Size
drum YES	FIRST CATEGORY:	Entity	Creation	Structure	
urum	arum YES	SECOND CATEGORY:	Entity	Creation	Artifact
sieve	NO	FIRST CATEGORY:	Entity	Creation	Artifact
sieving	sieving YES	FIRST CATEGORY:	Process	Elimination	
sieving	163	SECOND CATEGORY:	Process	Phase	Phase of treatment
tank	YES	FIRST CATEGORY:	Entity	Creation	Structure
talik	123	SECOND CATEGORY:	Entity	Creation	Artifact
aeration tank YES	YES	FIRST CATEGORY:	Entity	Creation	Structure
aeration tank	aeration tank YES	SECOND CATEGORY:	Entity	Creation	Artifact
calibration tank	YES	FIRST CATEGORY:	Entity	Creation	Structure
Cambradon tank	123	SECOND CATEGORY:	Entity	Creation	Artifact

Figure 4: Example of the categorization table.

From an ontological point of view, 16 categories were associated with attributes, 93 with entities, and 43 with processes. (For a full list of the conceptual category hierarchy in EcoLexicon and some examples of each category, see Appendix A.)

3. Ontological perspective in EcoLexicon

The ontological enhancement process in EcoLexicon was mainly based on the categorization of its concepts in semantic classes with a multidimensional approach. As a result, not only was it possible to improve the structuration and organization of all the environmental knowledge it contained, but also to offer new practical applications and functionalities so that the end user could make the most of the ontological information. Essentially, the ontologically-enhanced functions that were implemented in EcoLexicon are the following: (i) the ontological view, an optional addition to the conceptual networks displayed in the general view; and (ii) a new conceptual categories module, including the revised category hierarchy and a conceptual combinations function.

3.1 Ontological view

The general view of EcoLexicon includes a series of elements that show all the information contained in the database in a user-friendly interface that facilitates access to the different types of data. The main information about each entry is broken down into five modules: (i) definition module, with a terminological definition based on the explicitation of the *genus* and the *differentiae*; (ii) term module, with the lexical denominations for a concept in the different languages available and linguistic information; (iii) resource module, with multimodal resources such as images, videos and hyperlinks; (iv) conceptual categories module, with the list of categories to which the concept belongs; (v) phraseology module, with the phraseological pattern and the collocational information about the concepts and terms. Furthermore, this terminological knowledge base also offers more functionalities, including the possibility of searching specific concordances in the EcoLexicon corpus and extracting statistics about the information in the database.

The most prominent feature of EcoLexicon is its dynamic visual display of conceptual networks, where concepts are surrounded by their multilingual denominations and related to each other through semantic relations. In EcoLexicon, three different types of semantic relations are distinguished: generic-specific relations (type_of), part-whole relations (part_of, made_of, delimited_by, located_at, takes_place_in, phase_of), and non-hierarchical relations (affects, causes, attribute_of, opposite_of, studies, measures, represents, result_of, effected_by, has_function).

In relation to the ontological enhancement process in EcoLexicon, this visual display of conceptual networks was improved through the implementation of an optional feature known as the ontological view (Figure 5). As a result of the conceptual categorization, each concept in EcoLexicon is tagged with one or more of the 152 categories, which allows for including this information so that the end user can observe the combinatorial potential of concepts according to their ontological nature.

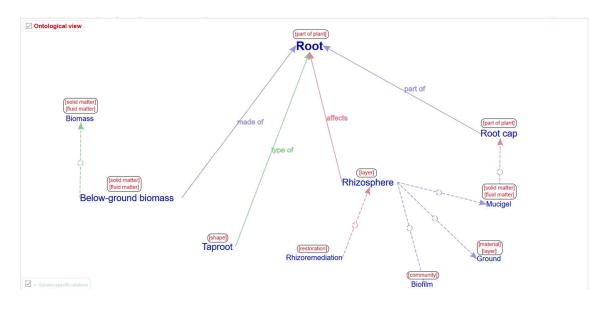


Figure 5: Ontological view (concept: root).

In Figure 6, the ontological view feature has been activated, so that a series of bubbles pop up over each concept (in blue) and indicate the conceptual categories to which each concept belongs (in red). Thanks to this functionality, there is a series of observations that can be made regarding the combinatorial potential of the chosen concept. For instance, it is interesting to confirm that solar cell (TRANSFORMING INSTRUMENT & PART OF INSTRUMENT) shares exactly the same categories with the other concepts to which it is related through a generic-specific relation: amorphous cell (TRANSFORMING INSTRUMENT & PART OF INSTRUMENT); crystalline solar cell (TRANSFORMING INSTRUMENT & PART OF INSTRUMENT); and thin-film solar cell (TRANSFORMING INSTRUMENT & PART OF INSTRUMENT). In the same way, since solar *cell* is categorized as a PART OF INSTRUMENT, its membership in larger conceptual categories \mathbf{is} expressed $\operatorname{through}$ part-whole relations: photovoltaic system (TRANSFORMING INSTRUMENT & SYSTEM) and solar panel (TRANSFORMING INSTRUMENT). Finally, the concepts that are linked to *solar cell* through

non-hierarchical relations are indeed related to the nature of this concept as a TRANSFORMING INSTRUMENT: *energy* (ENERGY & MEASUREMENT) and *solar radiation* (ENERGY MOVEMENT).

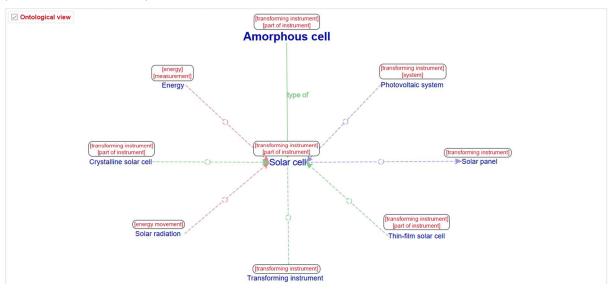


Figure 6: Ontological view (concept: *amorphous cell*).

3.2 Conceptual categories module

The original conceptual categories module in EcoLexicon only classified concepts according to the semantic roles designated in the Environmental Event (Faber, 2015; León-Araúz et al., 2012). For this reason, after performing the conceptual categorization process it was necessary to redesign this module. This involved two major changes: (i) the modification and update of the category hierarchy function; and (ii) the implementation of the conceptual combinations function. Figure 7 shows the conceptual categories module when selecting the concept *port*. Four conceptual categories (E-1.3: STRUCTURE, E-4.1: ARTIFICIAL GEOGRAPHIC FEATURE, E-4.2: NATURAL GEOGRAPHIC FEATURE, and E-12.1.2: FACILITY) are showcased, as well as the buttons for category hierarchy and conceptual combinations.



Figure 7: Conceptual categories module (concept: *port*).

3.2.1 Category hierarchy

The enhanced conceptual category hierarchy function of this new module contains a hierarchically-organized list of all 152 semantic classes (for a full list of the conceptual categories, see Appendix A). The members of each category can be accessed by clicking on the triangle to the left, enlarging the list to view the more specific subcategories (Figure 8). When a category is selected, a new window pops up with all the concepts belonging to it. This provides easy access to each entry, its information, and its ontologically-interrelated concepts in EcoLexicon (Figure 9). For example, in Figure 9 the concepts belonging to the DEFENSE STRUCTURE category are listed alphabetically, and clicking on any of them (e.g. *cofferdam*, *dike*) would lead EcoLexicon to its full entry with all the information.

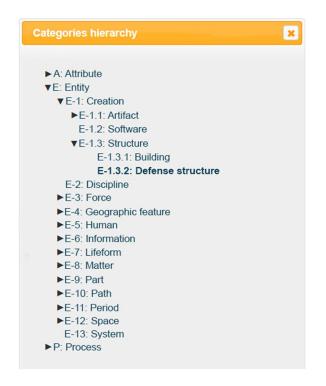


Figure 8: Category hierarchy function (category: DEFENSE STRUCTURE).

	Category E-1.3.2: Defense structure	×
Categories hierarchy	adjustable groin	^
 A: Attribute E: Entity E: 1. Creation E: 1. Artifact E: 1. 2: Software E: 1. 3: Structure E: 1. 3: Structure E: 1. 3: Structure E: 1. 3: Content of the structure E: 2: Discipline E: 3: Force E: 4: Geographic feature E: 5: Human E: 6: Information E: 7: Lifeform E: 8: Matter E: 9: Part E: 10: Path E: 10: Path E: 11: Period E: 2: Space E: 3: System 	apron arch dam barrier bed protection berm breakwater breakwater bulkhead buttress dam coastal defense cofferdam composite breakwater containment dike dam dike dam dike dam dike floating breakwater floating dock gabion	
	gravity dam	~

Figure 9: Category hierarchy function with examples (category: DEFENSE STRUCTURE).

3.2.2 Conceptual combinations

In the conceptual combinations function of the new conceptual categories module, users can perform a simple or advanced query. Figure 10 shows the query screen and the results screen of the simple query "hard structure". The simple query box can be used to perform a proximity search, since it then autocompletes with the available concepts as the user writes different letters. As shown in the results screen, the system automatically converts the user's search into a query expression ("hard structure [CONCEPT]") and displays a list of results in EcoLexicon that shows the combinatorial potential of the queried concept with other concepts through specific semantic relations. These results are, by default, collected under conceptual propositions made of conceptual categories (in black) linked through semantic relations (in orange). For instance, the fourth result in Figure 10 is listed as "[Defense structure] made of [Material]", but in order to see the specific concepts codified under those categories, it is necessary to click on the "+ Show specific results" option (in blue) next to this conceptual proposition, and thus the actual results of the query will appear: "HARD STRUCTURE made of CONCRETE", "HARD STRUCTURE made of STEEL", "HARD STRUCTURE made of QUARRY STONE", etc.

Conceptual combinations	Conceptual combinations
Query type: simple advanced Simple query: HARD STRUCTURE	Query: hard structure [CONCEPT] Query results:
Advanced query: + [concept] [semantic relation] [conceptual category]	1. [Defense structure] affects [Landform] + Show specific results: HARD STRUCTURE affects BOUNDED COAST 2. [Defense structure] has function [Energy movement] + Show specific results: HARD STRUCTURE has function [REFLECTION 3. [Defense structure] has function [Protection] + Show specific results: HARD STRUCTURE has function ABSORB WAVE ENERGY 4. [Defense structure] made of [Material] + Show specific results: HARD STRUCTURE made of CONCRETE HARD STRUCTURE made of CONCRETE HARD STRUCTURE made of QUARRY STONE
Search	< Back

Figure 10: Simple query (left side) and results (right side) in the conceptual combinations function using the expression "hard structure [CONCEPT]"

On the other hand, the advanced query presents a series of particularities that allow users to perform more complicated searches. As shown in Figure 11, the advanced query is based on three elements: (i) concepts; (ii) semantic relations; (iii) conceptual categories. By clicking on the orange bubbles next to the "+" symbol, users can add as many elements to the query as they want in any order, since this query allows for free element combination (e.g. "category + relation", "concept + relation + category", "category + relation + category", etc.). Similarly, any element can also be deleted. The concept bubble has a free text box to type anything, whilst the semantic relation and the conceptual category bubbles display a picklist showing all the relations or categories contained in EcoLexicon. However, it is also possible to choose the option "ANY" in the semantic relation and conceptual categories bubbles. In fact, displaying all the possibilities with a picklist is the simplest way for users to find and choose the most suitable option for their query. In addition, each bubble contains "AND" and "OR" buttons, which are useful if users want to look for more than one concept, relation and/or category found in the same position.

Figure 12 shows the query screen and the results screen of the advanced query "Water movement [CATEGORY] + any [SEMANTIC RELATION] + Natural water body [CATEGORY]". In order to perform this search, users must select the option "advanced" next to "Query type", and this will activate the advanced query box, where the user will then create a conceptual category bubble in order to select "Water movement", a semantic relation bubble in order to select "ANY", and a conceptual category bubble in order to select "Natural water body". As a consequence, this expression displays a series of results that include conceptual propositions linking concepts belonging to the WATER MOVEMENT category and the NATURAL WATER BODY category through any semantic relation. For instance, the first case is the conceptual proposition "[Water movement] affects [Natural water body]", including examples such as "FLOOD CURRENT affects BAY", "TIDE affects TIDAL RIVER", and "REGRESSION affects SEA".

mple query:		
I concept [semantic relation] [conceptual category] I OR 2 semantic relation I Concept Image: AND OR Image: AND OR I Concept Image: AND OR Image: AND OR I ND OR Image: AND OR Image: AND OR	3 conceptual category X AND OR	

Figure 11: Advanced query in the conceptual combinations function

Conceptual combinations 🔹	Conceptual combinations
Query type: simple advanced Simple query:	Query: Water movement [CATEGORY] + any [SEMANTIC RELATION] + Natural water body [CATEGORY] Query results:
Advanced query: + [conceptual category] 1 conceptual category & Water movement NY OR NY OR NY OR NY OR NY OR NY OR NY OR NY OR NY OR NY OR	1. [Water movement] affects [Natural water body] + Show specific results: FLOOD CURRENT affects SAY FLOOD CURRENT affects SETUARY OCEAN GYRE affects OCEAN FLOOD CURRENT affects ESTUARY TDEC affects IDAL RNER FLOOD CURRENT affects ESTUARY REGRESSION affects SAA SECHE affects LAKE 2. [Water movement] relation f [Natural water body] + Show specific results: INFLOW result of SPRING WAVE result of RIFFLE 3. [Water movement] relation [Natural water body] + Show specific results:
Search	< Back

Figure 12: Advanced query (left side) and results (right side) in the conceptual combinations function using the expression "Water movement [CATEGORY] + any [SEMANTIC RELATION] + Natural water body [CATEGORY]"

4. Conclusion

Contemporary theories of cognition have greatly influenced the most recent approaches to linguistics and terminology. Since terms are linguistic units that convey conceptual information dependent on the context, they cannot be analyzed in isolation, but rather as part of a situated environment where different brain modal systems interact. In the specific case of the development of terminological resources, it is essential to focus on how concepts are represented and organized in the mind or, in other words, on how conceptual information is categorized.

In addition, the influence of cognition on terminology has led to an enhancement of the ontological information displayed in linguistic and terminological resources, since it is necessary to portray more accurate representations of concepts and their information. Accordingly, more expressive formal ontologies benefit both human and non-human users by facilitating knowledge acquisition and offering a higher degree of interoperability, respectively. In this sense, EcoLexicon has experienced a process of ontological knowledge enhancement, mainly based on the categorization of its 4,500 concepts in 152 semantic categories. Thus, these top-down semantic categories distributed in up to five categorization levels were established to determine degrees of specificity and conceptual similarity, so that every concept could be tagged with a category showing its interrelation with other ontologically-related concepts.

As a result, not only it was possible to improve the structure and organization of the environmental knowledge contained in EcoLexicon, but also to offer new conceptual applications and functionalities, which benefitted from the ontological information that was implemented. Two new features derived from the conceptual categorization process were put in place: (i) the ontological view, an optional enhancement to the conceptual networks displayed in the general view that shows the combinatorial potential of concepts; and (ii) a revised conceptual categories module, including the modification and update of the category hierarchy function, and the inclusion of a new conceptual combinations function. This last feature is particularly useful for end users, since it allows them to perform simple and advanced queries regarding specific combinations of conceptual propositions (focusing on concepts, conceptual categories, and semantic relations).

In conclusion, this process of ontological enhancement in EcoLexicon will be useful not only for the improvements presented here in relation to the conceptual categories module, but also for the development of complementary features, such as the new phraseological module. More specifically, this last module would benefit from the integration of the category hierarchy into its functionalities, since it would make it possible to analyse phraseological units from an ontological approach.

Further research would require a series of users (experts and non-experts) to assess the main ontological features presented in this paper so as to validate their actual usefulness. Finally, since the future is based on interoperability among resources, it will be necessary to explore how the conceptual categorization can be implemented in the resources derived from EcoLexicon: the EcoLexicon corpus and EcoLexiCAT. Therefore, we plan to implement category annotation to enrich the EcoLexicon corpus, and ontological information derived from the conceptual categories module will be displayed in the EcoLexiCAT interface. Future work will also focus on how the ontological knowledge in EcoLexicon can be shared with external resources through Linked Data (León-Araúz et al., 2011a; León-Araúz et al., 2011b).

5. Acknowledgements

This research was carried out as part of project FFI2017-89127-P, Translation-Oriented Terminology Tools for Environmental Texts (TOTEM), funded by the Spanish Ministry of Economy and Competitiveness. Funding was also provided by the FPU grant given by the Spanish Ministry of Education and Professional Training to the first author of the article (ref. FPU16/02194).

6. References

- Barsalou, L. W. (2008). Grounded Cognition. Annual Review of Psychology, 59(1), pp. 617–645.
- Barsalou, L. W. (2010). Grounded Cognition: Past, Present, and Future. *Topics in Cognitive Science*, 2(4), pp. 716–724.
- Buttigieg, P. L., Morrison, N., Smith, B., Mungall, C. J. & Lewis, S. E. (2013). The Environment Ontology: Contextualising Biological and Biomedical Entities. *Journal of Biomedical Semantics*, 4(43).
- Buttigieg, P. L., Pafilis, E., Lewis, S. E., Schildhauer, M. P., Walls, R. L. & Mungall, C.J. (2016). The Environment Ontology in 2016: Bridging Domains with Increased Scope, Semantic Density, and Interoperation. *Journal of Biomedical Semantics*, 7(57).
- Evans, V. & Green, M. (2006). *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Faber, P. (2015). Frames as a Framework for Terminology. In H. J. Kockaert & F. Steurs (eds.) *Handbook of Terminology*, 1. Amsterdam/Philadelphia: John Benjamins, pp. 14–33.
- Faber, P. & San Martín, A. (2010). Conceptual Modeling in Specialized Knowledge Resources. Information Technologies and Knowledge, 4(2), pp. 110–121.
- Faber, P., León-Araúz, P. & Prieto-Velasco, J. A. (2009). Semantic Relations, Dynamicity, and Terminological Knowledge Bases. *Current Issues in Language Studies*, 1(1), pp. 1–23.
- Faber, P., León-Araúz, P. & Reimerink, A. (2014). Representing Environmental Knowledge in EcoLexicon. In E. Bárcena, T. Read & J. Arús (eds.) Languages for Specific Purposes in the Digital Era, Educational Linguistics, 19. Cham: Springer, pp. 267–301.
- Faber, P., León-Araúz, P. & Reimerink, A. (2016). EcoLexicon: New Features and Challenges. In I. Kernerman, I. Kosem, S. Krek & L. Trap-Jensen (eds.) Proceedings of GLOBALEX 2016: Lexicographic Resources for Human Language Technology in conjunction with the 10th edition of the Language Resources and Evaluation Conference. Portorož: ELRA, pp. 73–80.
- Gallese, V. & Lakoff, G. (2005). The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge. *Cognitive Neuropsychology*, 22(3/4), pp. 455– 479.
- Hahn, U. & Chater, N. (1997). Concepts and Similarity. In K. Lamberts & D. Shanks (eds.) Knowledge, Concepts, and Categories. Cambridge (MA)/London: MIT Press, pp. 93–131.
- Kageura, K. (1997). Multifaceted/Multidimensional Concept Systems. In S.E. Wright

& G. Budin (eds.) Handbook of Terminology Management: Basic Aspects of Terminology Management. Amsterdam/Philadelphia: John Benjamins, pp. 119–132.

- Kiefer, M., & Barsalou, L.W. (2013). Grounding the Human Conceptual System in Perception, Action, and Internal States. In W. Prinz, M. Beisert & A. Herwig (eds.) Action Science: Foundations of an Emerging Discipline. Cambridge (MA)/London: MIT Press, pp. 381–407.
- Kiefer, M. & Pulvermüller, F. (2012). Conceptual Representations in Mind and Brain: Theoretical Developments, Current Evidence and Future Directions. *Cortex*, 48(7), pp. 805–825.
- León-Araúz, P., Faber, P. & Magaña-Redondo, P. J. (2011a). Linking Domain-Specific Knowledge to Encyclopedic Knowledge: An Initial Approach to Linked Data. In Proceedings of the 2nd Workshop on the Multilingual Semantic Web (The 10th International Semantic Web Conference). Bonn, Germany, pp. 68–73.
- León-Araúz, P., Faber, P. & Montero-Martínez, S. (2012). Specialized Language Semantics. In P. Faber (ed.) A Cognitive Linguistics View of Terminology and Specialized Language. Berlin: De Gruyter Mouton, pp. 95–176.
- León-Araúz, P., Magaña-Redondo, P. J. & Faber, P. (2011b). Integrating Environment into the Linked Data Cloud. In W. Pillman, S. Schade & P. Smits (eds.) Proceedings of the 25th International Conference Environmental Informatics (Enviroinfo Ispra 2011). Aachen: Shaker Verlag, pp. 370–379.
- Mahesh, K. & Nirenburg, S. (1995). A Situated Ontology for Practical NLP. In Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, International Joint Conference on Artificial Intelligence (IJCAI-1995). Montreal, Canada, pp. 1–10.
- Mahon, B.Z. & Caramazza, A. (2009). Concepts and Categories: A Cognitive Neuropsychological Perspective. Annual Review of Psychology, 60(1), pp. 27–51.
- Martin, A. (2007). The Representation of Object Concepts in the Brain. Annual Review of Psychology, 58, pp. 25–45.
- Meteyard, L., Rodríguez Cuadrado, S., Bahrami, B. & Vigliocco, G. (2012). Coming of Age: A Review of Embodiment and the Neuroscience of Semantics. *Cortex*, 48(7), pp. 788–804.
- Moreno-Ortiz, A. & Pérez-Hernández, C. (2000). Reusing the Mikrokosmos Ontology for Concept-based Multilingual Terminology Databases. In Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC 2000). Athens, Greece, pp. 1061–1067.
- Murphy, G. L. & Lassaline, M. E. (1997). Hierarchical Structure in Concepts and Basic Level of Categorization. In K. Lamberts & D. Shanks (eds.) Knowledge, Concepts, and Categories. Cambridge (MA)/London: MIT Press, pp. 93–131.
- Weigand, H. (1997). Multilingual Ontology-Based Lexicon for News Filtering The TREVI Project. In K. Mahesh (ed.) Ontologies and Multilingual NLP: Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI). Nagoya, Japan, pp. 138–159.

Appendix A: Full conceptual category hierarchy in EcoLexicon

A: Attribute

- A-1: Ability [ex. autotrophic, permeability, tsunamigenic]
- A-2: Direction [ex. downstream, windward, onshore]
- A-3: Location [ex. hadopelagic, mesotidal, subaqueous]
- A-4: Measurement [ex. quantity, specific heat capacity, nutrient concentration]

A-4.1: Magnitude [ex. altitude, radicular zone depth, ambient temperature]

A-4.1.1: Level [ex. maximum flow, highest astronomical tide, freezing point]

A-4.1.1.1: Mean [ex. mean flow, mean tide level, average precipitation]

A-5: Origin [ex. Artificial, Aeolian, Lithologic]

A-6: Physical attribute [ex. color, soil texture, xericity]

A-6.1: Composition [ex. BIOCLASTIC, WOODY, MONOLITHIC]

A-6.2: Shape [ex. bacciform, euhedral, hook-shaped]

A-6.3: Size [ex. big, small, grain size]

A-6.4: State [ex. carbonate equilibrium, slope instability, unconsolidated]

A-6.4.1: Climate [ex. bioclimate, savanna climate, periglacialism]

A-7: Time [ex. Aperiodic, semidiurnal, temporary]

E: Entity

E-1: Creation [ex. wind turbine generator system, collector, septic system]

E-1.1: Artifact [ex. culvert, DC bus, STATOSCOPE]

E-1.1.1: Conduit [ex. drainage ditch, pipeline, duct]

E-1.1.2: Container [ex. cloud chamber, sedimentation tank, retention basin]

E-1.1.3: Instrument [ex. centrifugal pump, fishing net, weather satellite]

- E-1.1.3.1: Measuring instrument [ex. accelerometer, barometer, sounding machine]
- E-1.1.3.2: Recording instrument [ex. albedograph, marigraph, water-level recorder]
- E-1.1.3.3: Sampling instrument [ex. collector, automatic sampler, van Dorn Bottle]
- E-1.1.3.4: Transforming instrument [ex. upwind turbine, converter, solar cell]

E-1.1.4: Vehicle [ex. boat, dredger, electric vehicle]

E-1.2: Software [ex. computer application, Contour gridder, MODFLOW]

E-1.3: Structure [ex. spillway, pier, engineering structure]

 $E-1.3.1: Building \ [ex. Geothermal power plant, tide station, oil refinery]$

E-1.3.2: Defense structure [ex. reef breakwater, high groyne, retaining wall]

E-2: Discipline [ex. bioclimatology, human ecology, phytopathology]

E-3: Force [ex. tractive force, tectonic force, gravity]

E-3.1: Dynamics [ex. atmospheric dynamics, slope dynamics, coastal dynamics]

E-3.2: Energy [ex. electricity, wind energy, solar energy]

E-3.3: Stress [ex. friction, dynamic pressure, tension]

E-4: Geographic feature [ex. entry channel, aquifer, biome]

E-4.1: Artificial geographic feature [ex. groyne bay, quarry, port]

E-4.1.1: Artificial water body [ex. pool, pond, reservoir]

E-4.2: Natural geographic feature [ex. abyss, high plateau, bay]

E-4.2.1: Landform [ex. fan delta, river gorge, emergent coast]

E-4.2.1.1: Natural water body [ex. sea channel, Karst Spring, Lagoon]

- E-4.2.2: Landscape [ex. tidal shoal, monsoon forest, mangrove swamp]
- $E-5: Human \ [ex. \ port \ authority, \ human \ being, \ social \ agent]$
 - E-5.1: Institution [ex. meteorological service, city council, public institution]

E-5.2: Specialist [ex. geographer, geologist, oceanographer]

- E-6: Information [ex. piece of data, cartographic information, hydrologic data]
 - E-6.1: Classification [ex. climate classification, clade, urban hierarchy]
 - E-6.1.1: Scale [ex. Beaufort scale, state-of-sea scale, spectrum]
 - E-6.2: Document [ex. plan, protocol, tide table]
 - E-6.2.1: Law [ex. legislation, wildlife law, principle of environmental law]
 - E-6.3: Parameter [ex. structural criterion, quality indicator, K factor]
 - E-6.4: Record [ex. baseline cartography, echogram, meteorological series]
 - E-6.5: Representation [ex. geodatabase, auroral oval, soil profile]
 - E-6.5.1: Graph [ex. adiabatic chart, stratigraphic column, compound hydrograph]
 - E-6.5.2: Line [ex. rating curve, isohaline, meridian]
 - E-6.5.3: Map [ex. Nautical chart, orientation map, orthophotomap]
 - E-6.5.4: Mathematical expression [ex. coefficient, standard deviation, wave equation]
 - E-6.5.5: Model [ex. Ekman spiral, erosion model, simulation]
 - E-6.5.6: Picture [ex. Photomosaic, satellite image, orthophoto]
 - E-6.5.7: Unit [ex. steradian, faraday, milimeter]
 - E-6.6: Theory [ex. plate tectonics, equilibrium theory, stationary wave theory]
- E-7: Lifeform [ex. detritivore, native species, organism]
 - E-7.1: Animal [ex. Amphibian, Livestock, Crustacean]
 - E-7.2: Community [ex. benthos, biocenosis, biological community]
 - E-7.2.1: Animal community [ex. stygofauna, cohort, zooplankton]
 - E-7.2.2: Plant community [ex. phytobentos, flora, phytoplankton]
 - E-7.3: Fungus [ex. Basidiomycota, mycobiont, facultative parasite]
 - E-7.4: Microorganism [ex. bacteria, facultative aerobe, enteric virus]
 - E-7.5: Plant [ex. chamaephyte, phycobiont, mangrove]
- E-8: Matter [ex. greybody, organic material, substance]
 - E-8.1: Chemical substance [ex. carbonic acid, arsenic, nitrogen dioxide]
 - E-8.2: Fluid matter [ex. tar, lava flow, mud]
 - E-8.2.1: Fluid astronomical body [ex. heavenly body, star, Sun]
 - E-8.2.2: Gas [ex. polar Air, exhaust gas, smog]
 - E-8.2.3: Water [ex. runoff water, drinking water, rainwater]
 - E-8.2.3.1: Cloud [ex. altostratus, stratocumulus, frontal fog]
 - E-8.3: Particle [ex. volcanic ash, interleukin, ultrafine particle]
 - E-8.4: Solid matter [ex. solid fuel, solid waste, solute]

E-8.4.1: Deposit [ex. alluvium, sediment flow, aeolian deposit]

- E-8.4.2: Material [ex. CEMENT, REINFORCED CONCRETE, SEMICONDUCTOR]
 - E-8.4.2.1: Mineral [ex. anthracite, coarse sand, zeolite]
 - E-8.4.2.2: Rock [ex. limestone, quartz diorite, clastic sedimentary rock]
 - E-8.4.2.3: Soil [ex. leptosol, mollisol, saturated soil]

E-8.4.3: Snow/ice [ex. avalanche, snowflake, anchor ice]

E-8.4.4: Solid astronomical body [ex. Asteroid, planet, satellite]

E-9: Part [ex. discards, section, stator]

E-9.1: Part of instrument [ex. Anemometer mast, wind turbine rotor, flap]

E-9.2: Part of landform [ex. beach head, berm crest, soil properties]

E-9.3: Part of lifeform [ex. Allele, cell wall, tissue]

E-9.3.1: Part of animal [ex. Eosinophil, Otolith, Valve]

E-9.3.2: Part of fungus [ex. ascospore, sporocarp, paraplectenchyma]

E-9.3.3: Part of plant [ex. bracteole, chloroplast, dehiscent fruit]

E-9.4: Part of structure [ex. harbour mouth, spillway crest, groyne head]

E-9.5: Part of vehicle [ex. gunwale, hull, keel]

E-9.6: Part of water body [ex. downstream, Aphytal zone, sea floor]

E-10: Path [ex. road, gully, viaduct]

E-10.1: Imaginary path [ex. planetary orbit, ecliptic plane, Earth's elliptic orbit]

E-11: Period [ex. lunar day, autumn, useful life]

E-11.1: Era [ex. Devonian, Mesozoic Era, Pleistocene Epoch]

E-12: Space [ex. capillary interstice, medium, ecological niche]

E-12.1: Area [ex. sedimentary environment, protected area, ecoregion]

E-12.1.1: Administrative area [ex. city, municipal boundary, The United States of America]

E-12.1.2: Facility [ex. BIOMASS POWER PLANT, MEASURING STATION, GAUGING SITE]

E-12.1.3: Land [ex. basin slope, meadow, area of land]

E-12.2: Layer [ex. atmosphere, planetary boundary layer, lower mantle]

E-12.3: Limit [ex. wave crest, limit of uprush, amphidromic point]

E-12.4: Position [ex. BIFURCATION, DEPOCENTER, PERIGEE]

E-13: System [ex. detritus food chain, network, isolated system]

P: Process

P-1: Action [ex. biological action, spawning, environmental crime]

P-1.1: Analysis [ex. sedimentological analysis, environmental impact assessment, weather forecast]

P-1.2: Chemical reaction [ex. combustion, anabolism, deflagration]

P-1.3: Collection [ex. energy storage, soil water retention, sand trapping]

P-1.4: Interaction [ex. interspecific competition, air-sea interaction, endogenic geological process]

P-1.5: Management [ex. coastal management, sustainable water use, waste management]

P-1.6: Measurement [ex. stream gauging, densitometry, stochastic process]

P-1.7: Protection [ex. Absorb wave energy, soil conservation, flood prevention]

P-2: Activity [ex. subsistence agriculture, shifting cultivation, factory farming]

P-3: Addition [ex. tectonic accretion, artificial nourishment, phosphate fertilization]

P-4: Change [ex. climate change, ecological degradation, environmental impact]

P-4.1: Change in size/intensity [ex. tide acceleration, cyclogenesis, anticyclolysis]

P-4.1.1: Decrease [ex. retard littoral drift, wave setdown, reduction in longshore transport]

 $P-4.1.2: Increase \ [ex. sea level rise, algal bloom, rise of the water table]$

 $P-4.2: \ Change \ of \ direction \ [ex. deflection, density stratification, secular variation]$

P-4.3: Change of state [ex. condensation, soil liquefaction, solidification]

P-4.4: Disease [ex. bronchitis, yellow band disease, monilia disease]

P-4.5: Division [ex. cleavage, dispersion, breaking drops]

P-4.6: Transformation [ex. acidification, metamorphism, territorial transformation]

P-4.6.1: Pollution [ex. Atmospheric Pollution, ozone Pollution, ocean dumping]

P-4.6.2: Restoration [ex. bioremediation, environmental recovery, revegetation]

P-5: Cycle [ex. tidal cycle, carbon cycle, hydrologic cycle]

P-6: Elimination [ex. deforestation, mass extinction, elimination of solid waste]

P-7: Emission [ex. particulate emission, hydromagmatic eruption, evapotranspiration]

P-8: Formation [ex. Breccia formation, Atmospheric Ionization, Primary Production]

P-9: Loss [ex. coastal degradation, internal erosion, mechanical weathering]

P-10: Method [ex. air layering, hodograph method, polyculture]

P-11: Movement [ex. drift, osmosis, traffic]

P-11.1: Earth/soil movement [ex. continental drift, slope movement, tectonic earthquake]

P-11.2: Energy movement [ex. forced convection, atmospheric radiation, cloud electrification]

P-11.3: Fluid movement [ex. capillarity, laminar flow, magma intrusion]

P-11.3.1: Water movement [ex. coastal circulation, drift current, gravity flow]

P-11.4: Transport [ex. transfer, longshore transport, upwell]

P-11.5: Wave [ex. regular wave, atmospheric wave, progressive wave]

P-11.6: Wind movement [ex. SEA BREEZE, ANTICYCLONIC CIRCULATION, WARM FRONT]

P-12: Phase [ex. Karyokinesis, Cytokinesis, Preliminary Treatment]

P-12.1: Phase of cycle [ex. tidal stage, lithogenesis, orogeny]

P-12.2: Phase of treatment [ex. primary sedimentation, thermophilic digestion, preaeration]

P-13: Phenomenon [ex. lunar eclipse, extreme event, environmental noise]

P-13.1: Atmospheric phenomenon [ex. squall, advective thunderstorm, tropical cyclone]

P-13.1.1: Precipitation [ex. hydrometeor, freezing rain, convective precipitation]

P-13.2: Optical phenomenon [ex. rainbow, auroral storm, lightning flash]

This work is licensed under the Creative Commons Attribution ShareAlike 4.0 International License.

http://creativecommons.org/licenses/by-sa/4.0/

