Framing Terminology: A Process-Oriented Approach

Pamela Faber Benítez, Carlos Márquez Linares et Miguel Vega Expósito

Résumé de l’article

La notion de schéma utilisée dans la Sémantique des schémas est dérivée des grammaires des cas, qui étaient censées caractériser une petite situation abstraite de telle sorte que, si on souhaitait comprendre la structure sémantique d’un verbe, il était nécessaire de comprendre les propriétés de la scène entière qu’il activait. Un schéma a été défini plus amplement comme tout système de concepts mis en rapport de telle sorte qu’un concept évoque le système entier. Dans ce sens, il a une affinité évidente avec la terminologie, qui est également basée sur cette organisation conceptuelle. Néanmoins, bien que la Sémantique des schémas ait été appliquée à la lexicologie et à la syntaxe, jusqu’ici elle n’a pas été systématiquement appliquée à la terminologie. Ce document présente des arguments en faveur d’une organisation de domaines spécialisés s’appuyant sur des schémas, où un schéma dynamique et orienté sur les processus fournit le soutien conceptuel pour la localisation de sous-hiérarchies de concepts dans un événement de domaine spécialisé; l’élaboration d’un modèle de définition, permettant une représentation plus appropriée des domaines spécialisés et fournissant une meilleure façon de lier les termes avec les concepts. Le domaine de l’ingénierie côtière est utilisé comme exemple parce que les entités en jeu participent dans des processus qui sont difficiles à décrire au seul moyen des arbres conceptuels. Par l’utilisation des données de corpus nous démontrons comment il est possible de représenter cet événement et de créer un cadre dynamique qui enrichit et améliore le sens de la compréhension des concepts des domaines spécialisés.
Framing Terminology: A Process-Oriented Approach

PAMELA FABER BENÍTEZ
University of Granada, Granada, Spain
pfaber@ugr.es

CARLOS MÁRQUEZ LINARES
University of Granada, Granada, Spain

MIGUEL VEGA EXPÓSITO
University of Granada, Granada, Spain

RESUME
La notion de schéma utilisée dans la Sémantique des schémas est dérivée des grammaires des cas, qui étaient censées caractériser une petite situation abstraite de telle sorte que, si on souhaitait comprendre la structure sémantique d'un verbe, il était nécessaire de comprendre les propriétés de la scène entière qu'il activait. Un schéma a été défini plus amplement comme tout système de concepts mis en rapport de telle sorte qu'un concept évoque le système entier. Dans ce sens, il a une affinité évidente avec la terminologie, qui est également basée sur cette organisation conceptuelle. Néanmoins, bien que la Sémantique des schémas ait été appliquée à la lexicologie et à la syntaxe, jusqu'ici elle n'a pas été systématiquement appliquée à la terminologie.

Ce document présente des arguments en faveur d’une organisation de domaines spécialisés s’appuyant sur des schémas, où un schéma dynamique et orienté sur les processus fournit le soutien conceptuel pour la localisation de sous-hiérarchies de concepts dans un événement de domaine spécialisé; l’élaboration d’un modèle de définition, permettant une représentation plus appropriée des domaines spécialisés et fournissant une meilleure façon de lier les termes avec les concepts. Le domaine de l’ingénierie côtière est utilisé comme exemple parce que les entités en jeu participent dans des processus qui sont difficiles à décrire au seul moyen des arbres conceptuels. Par l’utilisation des données de corpus nous démontrons comment il est possible de représenter cet événement et de créer un cadre dynamique qui enrichit et améliore le sens de la compréhension des concepts des domaines spécialisés.

ABSTRACT
The frame notion used in Frame Semantics can be traced to case frames, which were said to characterize a small abstract situation in such a way that if one wished to understand the semantic structure of a verb it was necessary to understand the properties of the entire scene that it activated. A frame has been more broadly defined as any system of concepts related in such a way that one concept evokes the entire system. In this sense, it bears an obvious affinity with terminology, which is also based on such conceptual organization. However, despite the fact that Frame Semantics has been usefully applied to lexicology and syntax, so far it has not been systematically applied to terminology.

This paper argues for a frame-based organization of specialized fields in which a dynamic process-oriented frame provides the conceptual underpinnings for the location of sub-hierarchies of concepts within a specialized domain event, and the elaboration of a definition template, thus opening the door to a more adequate representation of specialized fields as well as supplying a better way of linking terms to concepts. The domain of coastal engineering is used as an example because the entities in play take part in processes that are difficult to describe only by means of conceptual trees. Through the use of corpus data we demonstrate how it is possible to represent such an event and create a dynamic frame which enriches and enhances the understanding of specialized field concepts.
1. Introduction

The principled configuration of concepts is essential in terminology, which targets the understanding and acquisition of specialized knowledge. For the most part, mainstream terminology theory is based on semantic memory models of the human conceptual system (Smith 1978), which are taxonomically organized. Such static knowledge representations traditionally take the form of conceptual trees, the description of which can be found in virtually all terminology handbooks (Sager 1990, Cabrè 1993, Wright and Budin 1997).

However, as shall be seen, recent research in cognitive neuroscience shows that this type of static, a modal vision of conceptual structure is lacking in psychological adequacy. According to Barsalou (2003: 552), the conceptual system develops to serve situated action, and at the broad level of organization, an important class of categories arises to streamline the action-environment interface. Consequently, a more believable representational system, and one more in consonance with empirical evidence would thus be dynamic and process-oriented with goal-derived categories that provide mappings from roles in action sequences to instantiations in the environment.

All of this evidently involves elaborating a more complex representation of concepts in specialized domains, but it is something that those involved in terminology processing and management cannot ignore. Such a representation would necessarily have to take propositional inferences into account and relate categories within some type of general event structure. The linguistic theoretical frameworks most applicable here are those with an enriched semantic component, such as Construction Grammar (Goldberg 1995) or Frame Semantics (Fillmore 1985). This has obvious repercussions in the representation of specialized knowledge.

In the first section of this paper we briefly describe the situated simulation theory of categorization against the background of other theories of conceptual organization, and explain how Frame Semantics, on the basis of data extracted from corpus analysis, can be used to represent a specialized knowledge field as a dynamic construct made up of clusters of processes and events. The second section shows how such a representation can be applied to the domain of coastal engineering.

2. Categorization and Terminology

Terminology is not only a matter of terms and term entries that endeavour, no matter how imperfectly, to represent slices of objective reality. In this sense, the representation of a specialized field should be more than a list or even a configuration of objects linguistically translated into either simple or compound nominal forms. It is necessary to situate concepts in a particular setting and within the context of dynamic processes that define and describe the principal event in the specialized field in question. As we shall see, the representation of a conceptual system cannot be static, but concepts should rather be organized around an action-environment interface (Barsalou 2003: 513). In terminology processing this would signify the specification of prototypical actions, which would activate schemas relating a specific set of participants.

2.1 Situated Simulation Theory

Cognitive psychology has proposed four theories of the human conceptual system: semantic memory, exemplar models, feed-forward connectionist nets, and situated simulation theory. Barsalou (2003) compares these models on the basis of the following dimensions: (1) architecture (modular vs. non-modular); (2) representation (amodal vs. modal); (3) abstraction (decontextualized vs. situated); (4) stability (stable vs. dynamical); (5) organization (taxonomic vs. action-environment interface), and offers empirical evidence which makes a strong case for the existence of situated simulations as a powerful interface between cognition and perception. He argues that such simulations underlie a situated, dynamic conceptual system, which would be neither fully modular
nor fully amodal. This kind of situated conceptualisation is conceived as a package of situation-specific inferences, which would include four types of information: (1) contextually-relevant properties of the focal category; (2) information about the background setting; (3) likely actions that the agent could take to achieve an associated goal; (4) likely introspective states that the agent might have while interacting with the category.

Restrictions on time and space do not permit a description of all of the empirical evidence offered in favour of this type of model, but very concisely, its application to specialized language is relevant for the following reasons:

- Sensory-motor representations are crucial to language because mental models of text meaning have spatial qualities, and it can be said that they underlie comprehension ability both for pictured events and for texts. It is thus likely that such representations would occur in the conceptual system. For example, in the translation of specialized texts, it is important for the translator, who is not an expert in the domain, to possess a spatial representation or a mental model of the processes and concepts within the text in order to fully understand them.

- Rather than a detached database of encyclopaedic knowledge (e.g. conceptual trees), the conceptual system bears a greater resemblance to an agent-dependent instruction manual, tailoring conceptual representations to the needs of situated action. This is evident in the fact that across different situations a concept can generate different conceptualisations, each designed to optimise one particular type of situated action with the respective category.

In terminology management this explains the multi-dimensional representation advocated for concepts in specialized domains, and which has been highlighted by a number of researchers (Bowker and Meyer 1993; Bowker 1997). In order to reflect all dimensions of meaning, concepts must be situated within the context of an event specific to the specialized area. In this sense the notion of “frame” is a useful construct not only to define concepts (Strehlow 1997), but also to situate them within a context in which categories are related to each other (Faber 2004).

2.2 Frame Semantics

The frame notion used in Frame Semantics can be traced to case frames (Fillmore, 1976), which were said to characterize a small abstract situation in such a way that if one wished to understand the semantic structure of a verb it was necessary to understand the properties of the entire scene that it activated (Fillmore 1982:115). A frame has been more broadly defined as any system of concepts related in such a way that one concept evokes the entire system.

In building a frame network, classification is involved since these networks are divided into domains, the domains into frames, and the frames can go through several levels of specificity by using hierarchical inheritance. Data is extracted by means of corpus analysis to encode underlying propositional structure, and define semantic roles. The elements of a frame may be shared with other frames because a lexical object can have several meanings, or the same dictionary meaning may have different social (connotative) meanings across situations.

In this sense, it is pertinent to terminology, which is based on such conceptual organization as well. Frames are also related to situated simulation theory in that their purpose is to map action and configure conceptual meaning more effectively within the real world. However, despite the fact that Frame Semantics and corpus analysis has been usefully applied to general language in the FrameNet project (Fillmore and Atkins 1998), so far it has not been systematically applied to specialized language.

3. Event Structure in Terminology

As previously mentioned, specialized knowledge domains can and should be regarded as configurations of complex events. In a previous project on medical terminology, conceptual categories were initially configured within the context of a medical event to facilitate understanding (Faber 2003). In a research project presently being carried out in the domain of coastal engineering,
we have found it useful to construct a similar type of event since an area such as coastal management needs to be understood as process-oriented, and the terms that codify this knowledge must be interrelated within such a configuration as components of a dynamic process.

Terminological studies generally focus on object concepts, which in most cases are linguistically represented by nominal forms. However, both in the comprehension and structure of specialized discourse, verbs play a crucial role. This is due to the fact that a considerable part of our knowledge is composed of EVENTS and STATES, many of which are linguistically represented by verbs.

As semantic predicates, verbs generally determine the overall form and meaning of sentences, which are the linguistic representation of one or various propositions. Goldberg (1998: 205) calls basic linguistic representations “constructions”, and proposes the following hypothesis:

**Scene-encoding hypothesis**: constructions that correspond to basic simple sentence types encode as their central senses, event types that are basic to human experience.

She argues that such constructions serve to carve up the world into discretely classified event types. This is in line with Langacker (1991: 294-295) who affirms that language in general is structured around certain conceptual archetypes.

In the description of an event within an engineering context, certain predicates are activated more frequently than others. The semantic parameters of their arguments contribute to constrain the meaning of the predicate and limit the possibility of polysemy. For example, in coastal engineering there is frequent reference to the concept of BEACH NOURISHMENT as a soft engineering solution to replenish coasts and beaches. Thus, the verb “nourish”, as derived from concordance analysis, is used in a way that is more specific than its usual meaning in general language, where the arguments are considerably less specific.

<table>
<thead>
<tr>
<th>nourish (general language)</th>
<th>to give sb/sth what is needed in order to live, grow or stay healthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>nourish (coastal engineering)</td>
<td>to give sth [coast, beach, shoreline] what is needed [sediment, sand] in order for it to stay healthy</td>
</tr>
</tbody>
</table>

The definitions above show that whereas in general language one can “nourish” by giving practically any type of sustenance, both concrete and abstract (e.g. food, water, education, affection), in specialized domains things are quite different. The nature of the arguments (beach, sand, etc.) restricts the meaning of general language verbs when used in specialized discourse.

In general communication such verbs are often highly polysemic, their multiple related meanings being the result of metaphorical extension. It is often the case that in specialized texts, the meaning of such predicates is restricted to one meaning, generally the one where the semantic arguments all fall into a specific knowledge area. Metaphor may also be present in specialized texts, but at other levels (Faber and Márquez Linares 2005).

### 3.1 Corpus Analysis

Evidently, the best way to find such data concerning predicates and their meaning is to extract it from a large corpus of texts. Corpus studies eloquently show that when a context of a lexical item is made explicit, information regarding its meaning and use also comes to light (Pearson 1998: 191). According to L’Homme et al. (1999: 32-33), who have used corpus analysis in terminology, context helps to recover information about related terms, definitional features, synonyms, as well as explicit relations between semantically related concepts. Furthermore, on the basis of context it is possible to verify the conceptual and grammatical information extracted from reference books, specify the meaning of abbreviations, and also to verify encyclopaedic information to include in the definition (Faber, López Rodríguez and Tercedor Sánchez 2001).

In this sense, the use of corpus analysis programs which elaborate frequency lists, preprocess the text before analysis, or automatically generate concordances of a term is not sufficient. It is necessary to interpret data with a coherent and systematic framework of textually-based linguistic
analysis (Bourigaut and Slodzian 1999: 29; L’Homme et al. 1999: 35). In terminology such a framework should include the information extracted from texts through corpus analysis, the consultation of dictionaries and reference books, terminological databases, and expert validation (Bourigaut and Slodzian 1999: 29-30). This methodology would only be valid if it focused on the communicative and socio-cognitive nature of terms (Cabré 1999; Temmerman 2000).

The methodology that we use for our present project in coastal engineering is based on past research in corpus elaboration and management (Pérez Hernández 2002; López Rodríguez 2001), knowledge modelling based on ontology building (Moreno Ortiz 2000ab; Moreno Ortiz and Pérez Hernández 2000), the identification of category templates that reflect the internal organization of specialized domains (Faber, López Rodríguez and Tercedor 2001), the application of those templates to the representation of terminological information and the elaboration of definitions (Faber and Mairal 1999; Tercedor 1999; García de Quesada 2001), and the specification of events as a way of relating categories and configuring them in larger knowledge structures.

3.2 The Coastal Engineering Event

Within the field of coastal engineering there are a variety of important event types, all of which are focused on the protection of coastlines. Integrated coastal zone management requires balancing a wide range of ecological, social, cultural, governance and economic considerations. Integrated coastal zone management was defined at an International Coastal Zone Workshop in 1989 as "a dynamic process in which a coordinated strategy is developed and implemented for the allocation of environmental, socio-cultural, and sustainable multiple uses of the coastal zone." The key word here is “dynamic process”, which indicates the constantly changing nature of the interacting components.

The coastal zone can be divided into three main components: the sea, the beach, and the land behind the beach. The sea, or offshore area, extends from the low water mark seaward. The beach zone extends from the low water mark to the seaward edge of the coastal vegetation. The adjoining coastal land is the zone extending landward for some distance from the end of the beach.

The three subsystems described here interact in many ways, and the boundaries between them fluctuate. The coastal zone is a complex, highly productive environment, and the health of one ecosystem is intimately tied to the health of the other ecosystems in the area, and even to areas some distance away.

It is evident that the representation of such a domain necessarily involves the elaboration of a complex event in order to situate and contextualize the concepts within each subarea of the event structure. For the analysis of the frame structure of the Coastal Engineering Event (CEE) the evidence provided by the corpus specifically devised for the PuertoTerm project was taken as the basis for an inductive approach. The initial phase in this approach was the development of a comprehensive list of terminological units in the corpus. Once this inventory was available, the second phase was the creation of a frame structure in which all of them could fit in a coherent and systematic way. (See Appendix for the graphic representation of the resulting CEE).

The corpus data showed that the CEE is conceptualised as a process. This process is initiated by an agent; it affects a specific kind of patient; and it produces a result. These macrocategories (AGENT ➔ PROCESS ➔ PATIENT/RESULT) are the core organisational structure which encompasses the vast majority of terms in the domain, and provides an adequate model to represent their interrelationships. Additionally, there are more peripheral categories which include INSTRUMENTS that are typically used during the CEE, as well as a clearly distinguishable category where the concepts used for the measurement, analysis, and description of the processes taking place in the main event can be satisfactorily grouped together.

The structure of the AGENT, PROCESS and PATIENT/RESULT macrocategories are so complex that it is somewhat difficult to establish features which are common to all their members. Thus, both gales and builders are included in the AGENT category; erosion and managed retreat are instances of PROCESS, and flora and jetties can both be categorised as PATIENT/RESULT entities. Notwithstanding their dissimilarities, the relationships between these core categories are consistent and allow a simple, coherent explanation of the linguistic facts with discrete instruments. This
problem was solved by regarding the macrocategories as templates, thus allowing a discrete modular approach to the CEE in which the label “category” was used to group lexical units with a greater degree of internal consistency.

The AGENT and PROCESS templates present a neatly parallel structure. The relationship that can be established between entities which are NATURAL AGENTS and the NATURAL PROCESSES caused by them, is parallel to the relationship between HUMAN AGENTS and the ARTIFICIAL PROCESSES they originate. This parallelism extends to their subcategories, which in both cases include processes that involve movement, accretion or loss/substraction of materials. Naturally, the CONSTRUCTION subcategory is a typically human process that has no equivalent to natural processes. The fact that the same terms are used to refer to both natural and artificial processes (e. g. consolidation, recharge, nourishment) demonstrates the need for a common template. Including both categories in the same template has further advantages: firstly, it allows for the double membership of terms that refer to processes that can have either an artificial or a human agent; secondly, it reflects the fact that both types of processes affect the PATIENT/RESULT template in a similar way.

The PATIENT/RESULT category poses an interesting case. Coastal engineering deals with the highly dynamic processes that take place in the coastal environment. Beaches, cliffs, marshes, plants are affected by both natural and artificial processes or can even be the result of a natural or artificial process. Thus, a beach is by definition a sediment deposit and hence the RESULT of erosion (a natural process of loss), but it is at the same time AFFECTED by erosion.

This dynamic nature of natural processes can also be observed in artificial processes. Human processes RESULT in constructions and structures which become a part of the coastal environment and, as such, can be in turn AFFECTED by both natural processes and subsequent human processes. Therefore, the same entity which is categorised as the RESULT of a process in a text can be categorised as a PATIENT affected by a different process in another text. The best way to solve this is by including both AFFECTED and RESULT in the same template.

INSTRUMENT is a well defined category, which refers to the specific machinery and devices that are used as aids to human processes carried out in the coastal environment.

The fact that the bulk of the PuertoTerm corpus consists of academic texts (e.g. articles, manuals, specialised web sites) is the reason behind the DESCRIPTION template in the graphic representation of the CEE. Many of these studies dwell on the description of the CEE. In other words, they specifically highlight the entities affected and resulting from it, the processes that have an effect on it and the entities that originate these processes. For this reason, the DESCRIPTION template includes specialised lexical items that may refer to the qualities of any member of any of the categories in any of the templates. There are many examples of this:

- “wind pattern” is used to describe winds as NATURAL AGENTS;
- “littoral cell” is used to describe the morphology of the coast, which can be both the RESULT of or AFFECTED by a process;
- “transport rate” makes reference to the amount of sediment which is transferred as a result of a typically—but not exclusively—NATURAL PROCESS.

Science seeks to describe Nature in order to predict its course. It should not come as a surprise, then, that the quantity of terms related to measurement and categorisation of the CEE is so large. The importance of the DESCRIPTION template is underlined by the complexity and diversity of its subcategories, which range from the measurement and representation of the entities involved in the CEE to the specifically designed instruments that are used for their study or the terms that are used to predict their evolution and interrelationships.

From a more linguistic point of view, it should be noted that terminological units used to describe the CEE are mostly nouns. The verbs in the texts that have been studied are mainly variants of the non-specialised verbs which designate the way categories interact with each other: cause, carry out, affect, create, become. Nominalisation is a well-known phenomenon in scientific texts since it allows authors to detach themselves from the text, contributing to the objectivity typically claimed by science, and it is also the best way to express the description, connections, relationships
and attributes of processes. Processes become arguments in the syntagmatic structure and predicates are used to designate the way these processes are related to each other and specify their attributes.

The frame in the appendix is useful because it allows the possibility of inserting hierarchies of terms into a global framework. For example, a process such as BEACH NOURISHMENT, would be inserted within the frame slot of ARTIFICIAL PROCESS. It would be based on the corpus analysis of the argument structure of the concordances of verbs such as nourish, replenish, fill, feed, as well as their nominal forms. The analysis of these predicates, all of which belong to the lexical field of CAUSATIVE POSSESSION [TO CAUSE SB/STH TO HAVE STH] and thus have similar argument structure, gives the following mini-hierarchy with slots indicating horizontal relations, which can be inserted into the larger event within the category of ARTIFICIAL PROCESS.

<table>
<thead>
<tr>
<th>ARTIFICIAL PROCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEACH NOURISHMENT</td>
</tr>
<tr>
<td>LOCATION</td>
</tr>
<tr>
<td>long straight beach</td>
</tr>
<tr>
<td>adjacent to inlets</td>
</tr>
<tr>
<td>natural inlet with bypassing bar</td>
</tr>
<tr>
<td>deepened inlet without jetties</td>
</tr>
<tr>
<td>jetted inlet</td>
</tr>
<tr>
<td>pocket beach</td>
</tr>
<tr>
<td>seawalled beach</td>
</tr>
<tr>
<td>DESIGN PARAMETER</td>
</tr>
<tr>
<td>project length</td>
</tr>
<tr>
<td>wave height</td>
</tr>
<tr>
<td>erosion rate</td>
</tr>
<tr>
<td>sediment characteristics</td>
</tr>
<tr>
<td>PROCESS</td>
</tr>
<tr>
<td>location of sediment source</td>
</tr>
<tr>
<td>dredging</td>
</tr>
<tr>
<td>placement of nourishment material</td>
</tr>
<tr>
<td>renourishment</td>
</tr>
<tr>
<td>EFFECT</td>
</tr>
<tr>
<td>storm protection</td>
</tr>
<tr>
<td>recreational benefits</td>
</tr>
<tr>
<td>environmental benefits</td>
</tr>
<tr>
<td>increased wildlife habitat</td>
</tr>
<tr>
<td>increased beach flora habitat</td>
</tr>
</tbody>
</table>

The elaboration of hierarchical knowledge representations based on corpus data is a valid way of configuring concepts in specialized knowledge fields. However, such conceptual structures only acquire meaning when placed within the broader context of a frame like the one shown in the Appendix.

4. Conclusions

Traditionally, the approach to terminological structure has been a purely hierarchical one. Specific categories were defined as types of other more generic or superordinate categories. Although there is both linguistic and psycholinguistic evidence that supports the notion that concepts are categorised in terms of typological and partonymic relationships, it is also clear that categorisation is not limited to these types of relationships. Categorisation is an eminently dynamic process and knowledge representation should be able to reflect this dynamism.
This paper argues for a frame-based organization of specialized fields in which a dynamic process-oriented frame provides the conceptual underpinnings for the location of sub-hierarchies of concepts within a specialized domain event, and the elaboration of a definition template, thus opening the door to a more adequate representation of specialized fields as well as supplying a better way of linking terms to concepts. The domain of coastal engineering is used as an example because the entities in play take part in processes that are difficult to describe only by means of conceptual trees. Through the use of corpus data we demonstrate how it is possible to represent such an event and create a dynamic frame which enriches and enhances the understanding of specialized field concepts.

NOTE

1. This research is part of the project PUERTOTERM: knowledge representation and the generation of terminological resources within the domain of coastal engineering, BFF2003-04720, funded by the Spanish Ministry of Education.

REFERENCES


Appendix

AGENT TEMPLATE

natural agent
- Water
  Movement of water: waves, tide, currents
- Atmospheric
  Wind, storms,...
- Geological
  Earthquakes

human agent

CAUSES

PROCESS TEMPLATE

natural process
- Movement
- Accretion
- Loss

artificial process
- Construction
- Addition
- Subtraction
- Movement

AFECTS

DESCRIPTION TEMPLATE

Attributes of
- Speed
- Height

Measurement of
- Time
- Space

Representation of

Simulation of

Prediction of

Discipline for study of

Instruments for description of

Procedures of description of

PATIENT/RESULT TEMPLATE

patient
- Coast features
  Coastline
  River
  Seabed
  Island
- Water mass
- Material
- Fauna/flora

becomes

result
- Modified coastal area
- Material

AFFECTS

result
- Hard construction
  Part of
  Management of