Frames and Terminology:
Representing Predicative Terms in the Field of the Environment

Marie-Claude L'Homme **Benoît Robichaud**

Observatoire de linguistique Sens-Texte (OLST)
C.P. 6128, succ. Centre-ville
Montréal (Québec) H3C 3J7 CANADA

mc.lhomme@umontreal.ca benoit.robichaud@umontreal.ca

Abstract

Terminological resources have traditionally focused on terms referring to entities, thereby ignoring other important concepts (processes, events and properties) in specialized fields of knowledge. Consequently, large parts of the conceptual structure of these fields are not taken into consideration nor represented. In this article, we show how terms that refer to processes and events (and, to a lesser extent, properties) can be characterized using Frame Semantics (Fillmore, 1982) and the methodology developed within the FrameNet project (Ruppenhofer et al., 2010). More specifically, we applied the framework to a subset of terms in the field of the environment. Frames are unveiled first by comparing similarities between the argument structures of terms already recorded in a terminological database and the relationships they share with other terms. A comparison is also carried out with the lexical units recorded in FrameNet. Then, relations between frames are defined that allow us to build small conceptual scenarios that are specific to the field of the environment. These relations are determined on the basis of the set of relations listed in the FrameNet project. This article reports on the methodology, the frames defined up to now and two specific conceptual scenarios (**Risk_scenario** and **Managing_waste**).

1 Introduction

Traditionally, terminological resources have been designed as knowledge repositories and until recently the focus has been placed on finding ways to represent the knowledge conveyed by terms. In fact, in several terminological applications, terms are viewed as the linguistic components of knowledge structures (i.e. linguistic labels attached to nodes that represent concepts). This perspective has led to the design of domain ontologies (or less formal structures) in which concepts are linked via a network of relations (is-a, part-of, cause-effect, etc) and terms are disambiguated linguistic labels assigned to these concepts.

However, it has been pointed out that, although interesting, these knowledge structures have important drawbacks as far as linguistic aspects are concerned: 1. They tend to focus on terms that denote entities (expressed by nouns) and little consideration is given to processes and events; 2. Other types of units that could be relevant for terminology, such as predicative terms (that designate processes, events and properties) are not represented in a way that fully captures their meaning; 3. They either overlook the linguistic properties of terms altogether, or linguistic properties (such as variation) are taken into account in a peripheral component of the representation.

An increasing number of researchers proposed alternative methods to add linguistic components to terminological knowledge structures (Faber, 2006, 2012; Montiel et al., 2010, among others). Others have developed methods to describe terms as linguistic units with frameworks designed for the lexicon in general. An interesting aspect of this latter work is the consideration given to terms that have been overlooked in knowledge structures, i.e. predicative terms and more specifically verbs (Condamines 1993; Lerat 2002; L'Homme 1998; Lorente 2002).

It is generally recognized that both the relationship with knowledge and linguistic properties are important aspects of terminological description, and methods should be developed to merge them into resources. However, it seems that terminologists still struggle to find an adequate balance between conceptual and linguistic representations (L'Homme, 2014). One possible solution resides in frames or frame-like representations that attract the interest of an increasing number of researchers (Dolbey et al., 2006; Faber, 2006, 2012; Schmidt 2009, among others, see Section 3).

This work is licensed under a Creative Commons Attribution 4.0 International Licence. Page numbers and proceedings footer are added by the organisers. Licence details: http://creativecommons.org/licenses/by/4.0/

Zock/Rapp/Huang (eds.): Proceedings of the 4th Workshop on Cognitive Aspects of the Lexicon, pages 186–197,
Dublin, Ireland, August 23, 2014.

This is the solution we chose in this paper. More specifically, we applied principles based on Frame Semantics (Fillmore, 1982, 1985; Fillmore and Baker, 2010) and the methodology developed within the FrameNet project (Fillmore et al., 2003; Ruppenhofer et al., 2010) to linguistic data related to the field of the environment. A first part of this work was reported in L'Homme et al. (2014), in which frames were defined based on the contents of a resource containing environment terms (e.g., *change*, *impact*, *recycle*). In this paper, we summarize our methodology to discover frames, and report on what has been done to define relations between frames and build conceptual scenarios that represent processes and events in the field. We then describe two specific scenarios that apply to the field of the environment (**Risk_scenario** and **Managing_waste**).

2 Theoretical assumptions and motivations

Processes and events represent an important part of the set of concepts to be represented in many fields of knowledge. This is the case in environment where events (e.g., "storm", "melt", and "warming") and processes (e.g., "damage", "threaten") can be observed. However, traditional terminological models (and even less traditional ones, such as ontological representations) are not properly equipped to describe these concepts and account for their specific linguistic properties, namely the fact that they require arguments (*X changes Y*; *impact of X on Y*).

Frame Semantics (Fillmore, 1982; Fillmore and Baker, 2010) and its related application FrameNet (Ruppenhofer et al., 2010) are specifically adapted to account for these concepts and offer different means to represent their conceptual as well as their linguistic properties. Frame Semantics (FS) is based on the assumption that the meanings of lexical units (LUs) are constructed in relation to background knowledge, whose structure can be analyzed in terms of semantic frames. Frames are conceptual scenarios in which different participants (called *frame elements*, *FEs*) appear. For instance, the **Criminal_investigation** frame is defined as follows in FrameNet: This frame describes the process that involves the determination by an authority, the Investigator, of the circumstances surrounding an Incident by means of inquiry.

The frame states that there are three obligatory participants in this scenario (FEs): **Investigator**, **Incident**, and **Suspect** (other non-obligatory participants – *non-core FEs* – are also listed). Lexical units such as *clue.n, inquire.v, inquiry.n, investigate.v, investigation.n* evoke this frame. These lexical units and their participants are also annotated in selected sentences, thus linking the conceptual and linguistic representations levels of the description, as shown below for the verb *investigate*:

- NP police, sheriff, officer-T-(1)

  About 30 adults were arrested in raids on ten Children of God homes on Wednesday night by [$_{\text{Investigator}}$ police] **INVESTIGATING** $^{\textbf{Target}}$ [$_{\text{Incident}}$ claims of child abuse]. [$_{\text{Suspect}}$ INI]

  [$_{\text{Investigator}}$ Anti-terrorist officers] were quickly on the scene and **INVESTIGATING** $^{\textbf{Target}}$ [$_{\text{Incident}}$ the further reports of suspect devices]. [$_{\text{Suspect}}$ INI]

- T-NP allegation-(1)

  The Botswana government says that [$_{\text{Investigator}}$ it] will **INVESTIGATE** $^{\textbf{Target}}$ [$_{\text{Incident}}$ the torture allegations]. [$_{\text{Suspect}}$ INI]

- T-NP case-(1)

  [$_{\text{Investigator}}$ The union] is also **INVESTIGATING** $^{\textbf{Target}}$ [$_{\text{Incident}}$ a number of cases of child labour and the sexual abuse of children by employers]. [$_{\text{Suspect}}$ INI]

Frames can share relationships with other frames as shown in Figure 1. **Criminal_investigation** is a subframe of **Crime_scenario**, it is preceded by **Committing_crime** and precedes **Criminal_process**.

We believe that frames are well suited to represent the properties of predicative terms: annotations serve to capture their linguistic properties and link these properties to an abstract representation level, i.e. the frame. Furthermore, and this is what is explored in this paper, relations between frames, can help unveil larger conceptual scenarios in which these terms are involved. In FrameNet, some subject-specific frames can already be found, as shown below (Figure 1) with **Crime_scenario** and other related frames.

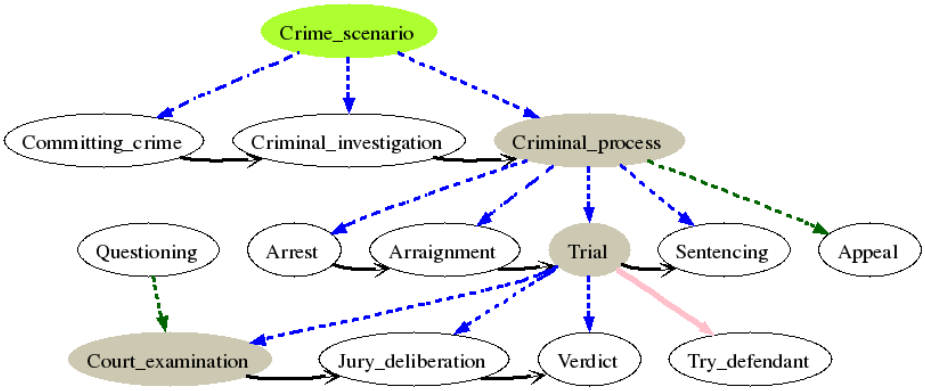


Figure 1. **Crime_scenario** and related frames in FrameNet (2014)

We assume that this can be applied to the frames of a specialized field such as the environment. However, we believe that the specialized lexicon will display some characteristics that will result in the necessity to define specific frames and perhaps specific scenarios. We explore this on a subset of data that is presented in Section 4.

3 Related work

In addition to projects aiming to describe the general lexicon in English (FrameNet, 2014), and in other languages, such as German, Japanese, and Spanish (Boas, 2009), an increasing number of researchers in terminology or related fields suggest that Frame Semantics (FS) or compatible frameworks are well suited to describe terms.

In Dolbey et al. (2006), Frame Semantics is adapted in order to develop frames in the field of biomedicine and link these frames to existing ontologies. Another application in medicine can be found in Wandji et al. (2013) where authors attempt to discover frames in the field with natural language processing techniques and an external resource (a medical terminology). Schmidt (2009) introduced some adaptations to the original framework of FS to account for multilingual data (English, French and German) in the field of soccer. Pimentel (2013) used the framework to establish equivalence relationships between English and Portuguese verbs in the field of law. L'Homme (2012) describes an annotation module added to two terminological resources (computing and environment) that is based on the annotation methodology developed within the FrameNet project. Finally, Faber (2012) refers to FS in order to account for concepts in the field of the environment and proposes a general frame (the environment event) to represent the interrelated processes and events observed in the field. The proposal has led to an approach in terminology called *Frame-based terminology*.

The work reported in this article bears some similarities with and differs from the work cited above in the following ways: 1. Contrary to some of this work, frames are discovered after terms are described rather than postulated prior to the descriptive work (we took a strictly bottom-up approach); 2. Frames are defined by observing similarities between terms (Sections 4.2 and 4.3); 3. Relations between frames are based on those already defined in FrameNet, but they must be valid from the point of view of the field of the environment. Hence, some differences are likely to be observed with similar frames appearing in FrameNet or with frames defined for other fields of knowledge.

4 Methodology

This section describes the data used in this work (extracted from a terminological database), and the different steps taken to unveil semantic frames and relations between them.

4.1 The terminological database

Our analysis is based on data recorded in an existing terminological database that contains terms in the field of the environment (it covers four subfields: climate change, residual material management,

electric transportation, and renewable energy).[1] The database – compiled chiefly according to the principles of Explanatory Combinatorial Lexicology (Mel'čuk et al., 1995) – contains terms in English, French, Portuguese and Spanish. Entries provide a description of the lexico-semantic properties of terms (Figure 2): actantial (i.e. argument) structure, linguistic realizations of actants (i.e. arguments), and lexical relationships (including paradigmatic relationships and collocations).

threat $_1$, n

Status : 2

a threat: ~ of ***change*** $_{1a}$ ⊛ to ***ecosystem*** $_1$ ⊛

Contexts
Lexical relations

Explanation	Related term
Related Meanings	
≈	hazard # $_1$
≈	risk $_1$
Opposites	
Opposite	preservation protection $_1$
Other Parts of Speech and Derivatives	
Verb	threaten $_1$
An ***ecosystem*** that has undergone a t.	threatened $_1$
Combinations	
The ***change*** causes a t.	pose a ~
A ***change*** causes that a t. be more important	increase a ~
The ***ecosystem*** undergoes a t.	experience a `~
Someone or something causes that a t. be less important	reduce a ~

Figure 2. Entry *threat* in the environment database

Terms recorded in the database are nouns (Eng. *biodiversity*, *energy*; Fr. *bioénergie*, *environnement*), verbs (Eng. *erode*, *pollute*; Fr. *électrifier*, *incinérer*), adjectives (Eng. *anthropogenic*, *global*; Fr. *aride*, *vert*) or adverbs (Eng. *globally, locally*). In this work, we took into account verbs, nouns (that refer to events or processes) and a small set of adjectives (e.g., Eng. *absorb*, *concentration*, *threatened*; Fr. *menacé*, *réchauffer*, *tri*). A subset of 169 English terms and 205 French terms underwent the analysis described in the following subsections.

4.2 Annotated contexts

In the database, several predicative terms (and all those that were selected for this analysis) come with up to 20 annotated sentences. The annotation is based on the methodology developed within the FrameNet project (Ruppenhofer et al., 2010). The original objectives of the annotations were twofold: 1. Show how actants (i.e. arguments) stated in the actantial (i.e. argument) structure are realized linguistically; 2. Supply terminologists writing entries with linguistic evidence to support their intuitions.

In annotations (Figure 3), the predicative unit appears in capital letters and in bold. Participants are divided into two different types: actants (in bold) correspond to obligatory participants (roughly equivalent to FN's core frame elements); circumstants are non-obligatory participants (that correspond roughly to non-core FEs). Participants appear in different colors according to their role (Cause, Patient, etc.). A table summarizes the different patterns found in annotations.

> The major challenges for the environment today are climate change, the decline in biodiversity, the **THREAT to our health from pollution**, the way in which we use natural resources and the production of too much waste. [CHANG_1EUROPAENV 0 TK MCLH 19/07/2012]
>
> Changes in frequency and intensity of extreme weather and **climate events** could pose a serious **THREAT to human health**. [CHANG_VULNERABILITY 0 TK MCLH 19/07/2012]
>
> The specific **THREAT to some of these ecosystems** is discussed in detail elsewhere in this paper. [CHANG_2IPCCBIODIVERSITE 0 TK MCLH 19/07/2012]

[1] The database is enriched on an ongoing basis. Hence, some terms can be added to frames already defined. Other subfields will also be taken into account in the future.

Population growth and **degradation of water quality** are significant **THREATS to water security in many parts of Africa**, and the combination of continued population increases and global warming impacts is likely to accentuate water scarcity in subhumid regions of Africa. [CHANG_3IPCCCONSEQUENCE 0 TK MCLH 19/07/2012]

Actants		
Cause	Complement (PP-from) Indirect link	degradation event growth pollution
Patient	Complement (PP-to)	ecosystem health security
Others		
Degree	Modifier (AP)	serious significant
Descriptor	Modifier (AP)	Specific
Location	Complement (PP-in)	Part

Figure 3. Annotations for the term *threat*

4.3 Identification of frames

In a previous study (L'Homme et al. 2014), we analyzed data contained in an environment database to establish whether some lexical units could be associated with frames similar to those that are recorded in FrameNet or potentially lead to new ones. The methodology for discovering frames consists basically in: 1. Extracting relevant data from the environment database; and 2. Using FrameNet data (in English) as a reference to identify a first set of existing frames that the terms in our database could evoke. A set of tools were devised to help us carry out the analysis.

Identifying similarities between terms encoded in the environment database

The first tool we use is a script that extracts relevant data from the English and French versions of the environment database and presents it in two separate sortable tables (where the sort function was programmed to fit specific criteria). These tables are helpful as they bring together, flatten and sort information that is normally distributed in different entries of the database. Along with the terms and their part of speech, the following information is presented in additional columns (Figure 4).

Terme	Act 1	Act 2	Act 3	Act 4	Circ	Frames	Liens
alter.1.en alter: Cause or Agent ~ Patient	Cause	Agent	Patient	_	Degree Manner	Cause_change	Caus@Degrad : environment.1.en Caus@Degrad : habitat.1.en Degrad@ : seasonality.1.en
degrade.1.en degrade: Cause or Agent ~ Patient	Cause	Agent	Patient	_	Degree Location Manner	Damaging	Caus@Degrad : environment.1.en Caus@Degrad : watershed.1.en CausDegrad : ecosystem.1.en
effect.1.en an effect: ~ of Cause or Agent on Patient	Cause	Agent	Patient	_	Degree Duration	Objective_influence	_
influence.1.en an influence: ~ of Cause or Agent on Patient	Cause	Agent	Patient	_	Descriptor Duration Time	Objective_influence	_
influence.2.en influence: Cause or Agent ~ Patient	Cause	Agent	Patient	_	Degree Duration	Objective_influence	_

Figure 4. Data presenting frame-relevant information for English terms

- Semantic roles of actants placed in four consecutive columns and in the order in which they appear in the actantial structure of the term entries;
- Semantic roles of circumstants extracted from the annotated contexts associated with the terms, ordered and displayed in a fifth column;
- A frame name (taken from an extra file used aside the database entries). This name was added once it was defined by the terminologist that carried out the analysis (see Section 4.3.4);

- Verbs and nouns associated with the LUs through specific collocation relationships in a last column.[2]

Identifying similarities between terms encoded in the DiCoEnviro

In addition to the tables described in Section 4.3.1, another script was written to present a comparison page that contains information related to terms from the environment database along with LUs recorded in FrameNet. Each English entry of the environment database is first searched in the last release of the FrameNet data (Baker and Hung, 2010)[3], and presented side by side with the corresponding lexical units from FrameNet when matches are found (Figure 5). More specifically, the script retrieves the following information:

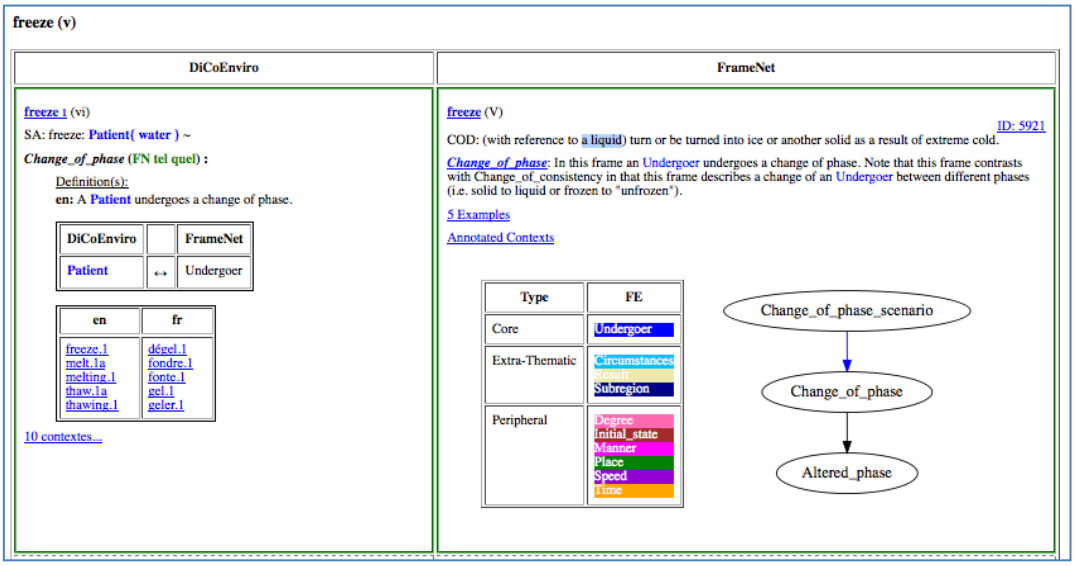


Figure 5. Comparison of environment terms with LUs in FrameNet

- From FrameNet: definitions of frames, their core and non-core FEs, relationships these frames have with other frames, and finally the annotated contexts accompanying the LUs themselves. A series of hyperlinks are also provided so that the terminologist analyzing the data can refer to FrameNet whenever necessary.
- From the environment database: actantial structures (i.e. showing the list of actants associated with the terms), the annotated contexts, and incidentally, for further stages of the analysis, the French and Spanish equivalents.

Differences between FrameNet and environment database

When comparing the data extracted from the environment database and FrameNet, we needed to take into consideration that the two resources bear some theoretical as well as methodological differences. We summarize them below:

- In FrameNet, FEs are defined at the level of frames while, in environment database, actants (and circumstants) are stated at the level of LUs. We established that terms in the environment database could evoke an existing frame if a relationship could be established between the set of core FEs and the actants, and if the FEs and actants were represented with comparable labels.

[2] Lexical relationships are represented in the database with lexical functions (LFs), a system developed in Explanatory Combinatorial Lexicology (Mel'čuk et al., 1995). In the online version, a natural language explanation is proposed (Figure 2): this explanation "translates" LFs' expressiveness in a way that is more accessible to users.

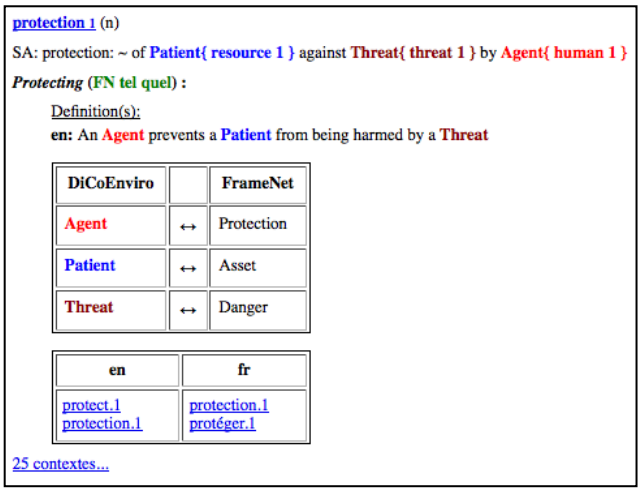
[3] For this, we used the XML files supplied by the FrameNet team. However, we noticed some differences with the online version of FrameNet: we needed to check whether the information had been updated.

- Secondly, due to the objectives of each resource, the number of core FEs in a frame could differ in comparison with the number of actants represented for a term in the environment database. Often, the number of core FEs was higher than the number of actants. In some cases, the environment database defines a participant as being a circumstant and a correspondence could be established with FrameNet. In other cases, the specificity of the specialized domain needed to be taken into consideration.
- Thirdly, labels used for most FEs are very specific since they are defined within a frame. In the environment database, labels are general and defined for the entire set of terms that are included in the database. In these cases, we generalized some of the labels. For example, labels such as Entity, Item, Theme, and Undergoer in FrameNet were assumed to correspond to Patient in the environment database.
- Fourthly, in FrameNet, different labels can account for an FE that would be realized in the same syntactic function. In the environment database, actants can be split (Agent or Cause for instance). In both cases, we considered these as being instantiations of the same argument position.

Assigning terms to frames

To make explicit the association of terms to frames (already recorded in FrameNet or especially created for the field of the environment), but also to facilitate the pairwise comparison of actants with FEs, we created an auxilary XML file aside from the files used to encode entries in the database (rather than adding this information in each terminological entry). Throughout the analysis, the file was enriched with additional information such as definitions and examples specific to the field of the environment, and relations frames have with other frames discovered or created (see Section 4.4).

Once created, the file can be loaded by the scripts mentioned earlier and used to help the analysis as it can be passed down to the comparison of terms and LUs. A comparison of actants and FEs is shown Figure 6.


protection 1 (n)

SA: protection: ~ of Patient{ resource 1 } against Threat{ threat 1 } by Agent{ human 1 }

Protecting (FN tel quel) :

Definition(s):

en: An Agent prevents a Patient from being harmed by a Threat

DiCoEnviro		FrameNet
Agent	↔	Protection
Patient	↔	Asset
Threat	↔	Danger

en	fr
protect.1 protection.1	protection.1 protéger.1

25 contextes...

Figure 6. Comparison of FEs in FrameNet and actants in the environment database

4.4 Frames discovered for environment terms

In L'Homme et al. (2014), we had analyzed 105 English and 159 French terms. This first set of data allowed us to find that some LUs were equivalent to frames already recorded in FrameNet; but that new frames also needed to be defined.

Currently, the different frames defined and the terms that evoke them appear in Table 1. The difference between English and French simply reflect the fact that more terms have been analyzed in French and in English up to now.

- **Entirely compatible**: The description of the terms in the environment database and the frames in FrameNet are similar (the number of actants vs. FEs and their semantic type is basically the same). For instance, *threaten* (Agent or Cause ~ Patient) evokes the **Endangering** frame (An Agent or Cause is responsible for placing a Valued_entity at risk).[4]

[4] Even if these frames are entirely compatible with those described in FrameNet, some differences are worth mentioning (in addition to those already taken into consideration when comparing the data, see Section 4.3.3). First, frames described in the environment database are much more restricted that those appearing in FrameNet

- **Alternation**: This category was created for cases where the environment database distinguishes two separate entries for closely related LUs. For instance, *predict*$_{1a}$ (Method ~ Patient) and *Predict*$_{1b}$ (Agent ~ Patient with Method) evoke a single frame, i.e. **Predicting** (An Agent states or makes known a Patient based on a Method; FN. A Speaker states or makes known a future Eventuality on the basis of some Evidence).[5]

Category	Number of Frames	Number of English LUs	Number of French LUs
Entirely compatible	19	45	52
Partly compatible	21	68	70
Alternation	2	8	9
New	31	60	85
Pending	6	9	13
TOTAL	79	190	229

Table 1: Different frames defined and number of LUs

- **Partly compatible**: The description of the terms in the environment database and the frames described in FrameNet are not exactly the same (the numbers of actants vs. frame elements differ). For instance, *risk* has three actants (~ of Result on Patient from Cause) and evokes the **Run_risk** frame, but the original frame has four core frame elements (Action, Asset, Bad_outcome, and Protagonist).
- **New**: Sets of new frames were defined for cases in which no existing frame could be found or cases where an existing frame was not well adapted for the environment. For instance, a new frame was created to LUs such as *recycle* and *recycling*, i.e. **Preparing_for_reuse**.
- **Pending**: Some LUs have been assigned to frames only provisionally for a number of reasons (few occurrences in the corpus, only one LU in the frame, etc.).

4.5 Identification of relations between frames

It soon became obvious that some frames defined for the field of the environment were related conceptually. We determined these relations using as a starting point the set of relations defined in the FrameNet project: these relations were sought in our data. We assumed that they would be valid – at least in part – for the domain of the environment, since they had been defined on a substantial amount of data.

This allowed us to discover conceptual scenarios specific to the field. In Table 2, we first describe the list of FrameNet relations taken into account and relations that are defined for the purpose of this project.

Relations used to link frames

The list of relations based on FrameNet are listed in Table 2.[6]

---

and the terms that evoke these frames may correspond to subsenses or microsenses (as defined by Cruse, 2011). For instance, the **Being_at_risk** frame in the environment applies only to things such as *species*, *ecosystems*, *plants*, etc. In addition, the number of terms that evoke a frame is often much lower than those recorded in FrameNet. For instance, the terms evoking the **Being_at_risk** frame in the environment data are the following: *sensitivity*, *threatened*, *vulnerability*, *vulnerable* (whereas in FrameNet, the list comprises: *danger.n, insecure.a, risk.n, safe.a, safety.n, secure.a, security.n, unsafe.a, vulnerability.n, vulnerable.a*).

[5] The alternation can be illustrated with the following examples: *Even the most sophisticated models cannot **predict** the details of how the climate change will unfold; it is also possible that our models will better enable us to **predict** the consequences.*

[6] Here, some differences with the way relations are defined in FrameNet are probably present. This part of the analysis is based on our interpretation of the way relations are defined in Ruppenhofer et al. (2010), the ones that appear in FrameNet and our own data.

Is causative of	This relation was established between the **Endangering** (with terms such as *endanger*, *threaten*) and the **Being_at_risk** (with terms such as *threatened* and *vulnerable*) frames. *Loss of important habitats (wetlands, tundra, isolated habitats) would THREATEN some species, including rare/endemic species and migratory birds.* *Low-lying island states and atolls are especially VULNERABLE to climate change and associated sea-level rise.*
Is inchoative of	This relation was defined between the **Cause_temperature_change** (with terms such as $cool_{1b}$, $warm_{1b}$) and **Change_of_temperature** (with terms such as $cool_{1a}$, $warm_{1a,}$, $warming_1$) frames. *… gases such as carbon dioxide (CO2) which WARM the Earth's surface.* *… the COOLING of the Northern Hemisphere may lead to increased warmth …*
Inherits from	This relation was established between **Change_position_on_a_scale** (with terms such as *decline*, *decrease*, *grow*, *increase*, *rise*) and **Change_of_temperature** (with terms such as $cool_{1a}$, $warm_{1a}$, *warming*). *The global average surface temperature has INCREASED over the 20th century by about 0.6 °C* *… the Earth will WARM in the near future.* The inverse relation is **Is inherited by.**
Has subframe	This relation was established between **Managing_waste** (with the terms *manage* and *management*) and the **Recover** (with terms such as Eng. *recover*, Fr. *récupérer*, *récupération*), **Removing** (with terms such as Eng. *discard*, *disposal*, Fr. *éliminer*), **Separating** (with terms such as *segregate*, *separate* and *sort*), and **Collecting** (with terms such as Eng. *collect*, Fr. *collecte*, *ramassage*) frames. *The Guelph wet-dry recycling centre can MANAGE up to 44,000 tonnes of compostables …* *… ensure that waste is RECOVERED or disposed of safely …* The inverse relation is **Is subframe of**.
Is perspectivized in	This relation was established between the **Greenhouse effect** (with the terms Eng. *greenhouse effect* and Fr. *effet de serre*) and the **Accumulating** (with terms such as Eng. *accumulate*, Fr. *accumulation*, *concentration*) and **Trapping** (with terms such as Eng. *trap*, Fr. *emprisonner*, *piéger*) frames. *The Earth has a natural GREENHOUSE EFFECT which keeps it much warmer that it would be without an atmosphere.* *If injected into the atmosphere, these gases ACCUMULATE there.* *… the atmosphere is slowly TRAPPING more heat over the years and enhancing the Earth's natural greenhouse effect.* The inverse relation is **Perspective on.**
Precedes	This relation was defined between the **Separating** and the **Removing** and **Recover** frames. *SEPARATE unwanted impurities and inorganic material …* *… all waste stabilization and DISPOSAL activities are preceded by some period of interim storage.* The inverse relation is **Is preceded by.**
Is used by	This relation was established between the **Protecting** (with Eng. *protect*, Fr. *protection*) and the **Run_risk** (with Eng. *risk* and Fr. *risque*) frames. *A greater number of people and those who are less indoctrinated seek to PROTECT humanity , even from itself …* *… the RISK of aggregate net damage due to climate change …* The inverse relation is **Uses**.

Table 2: Relations between frames of the field of environment

In addition to the relations based on those defined in the FrameNet project, we added new ones to capture some important conceptual perspective in the field of the environement:

- **Is opposed to**: This relation was established between the **Recover** and the **Removing** frames.
- Is a property of (has property): This relation was established between the **Judgment_of_intensity** (with LUs such as *intense*, *extreme*, *severe*) and the **Weather_event** frames (the latter one comprises LUs such as *event*, *activity*).

Up to now, among the 73 frames defined for the environment data, and about 70 are linked with one or two of the relations listed in this section. A small number of frames are linked provisionally with the **See also** relation. This simply indicates that a relation is present but its labelling is pending.

Displaying relations

After the creation of the auxiliary XML file recording the terms membership to frames described in Section 4.3.4, a search interface was designed and programmed to provide a more user-friendly access to its information.

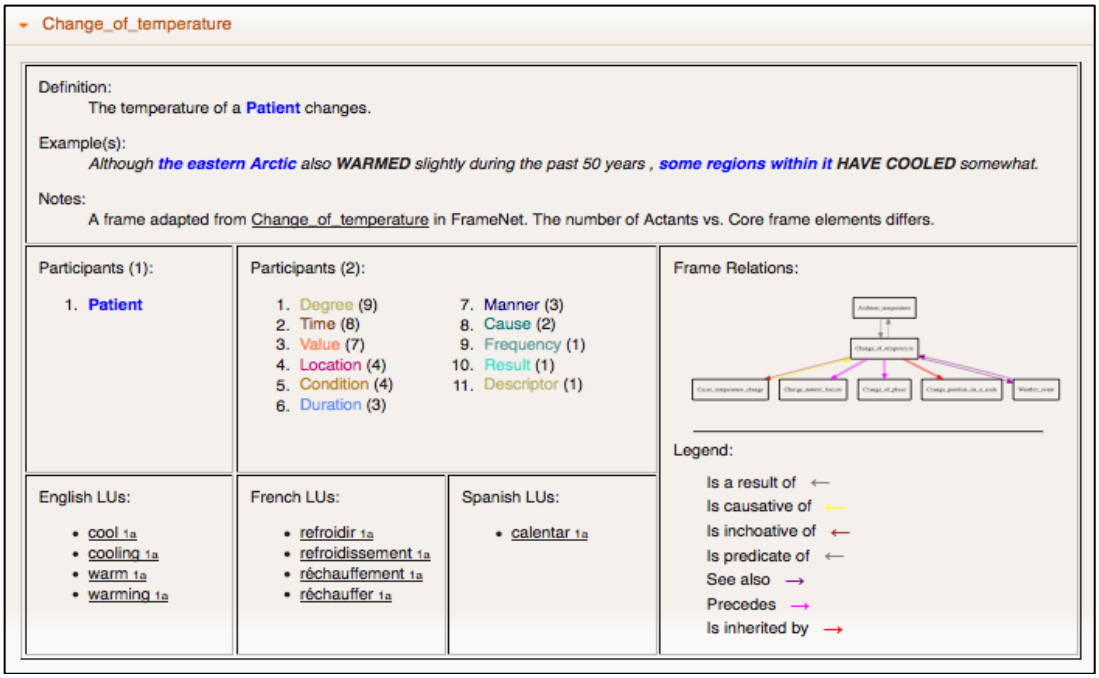


Figure 7: Display of the information associated to a frame in the field of environment

This interface allows us to select or search frames themselves, as well as terms or actantial roles. Search results display definitions, examples and notes associated to frames, their participants, together with lists of terms that evoke them. As in the FrameGrapher in FrameNet, rather than simply listing the relations that frames share with others frames, we present them as graphs (Figures 8 and 10). This provides a more comprehensive view of broader sets of frames and makes it easier to unveil some scenarios that we believe are specific to the field of the environment.

5 Two scenarios in the field of the environment

In this section, we describe two small conceptual scenarios that were discovered thanks to the establishment of relations described in Section 4. The first is the **Risk_scenario** that also appears in FrameNet. The second one is **Managing_waste** that has no direct counterpart in FrameNet (even though some frames appear to correspond to frames recorded in FrameNet). Other scenarios are in the process of being defined.

5.1 Risk_scenario

The **Risk_scenario** discovered on the basis of the data extracted from the environment database appears in Figure 8. We also reproduced the scenario proper to FrameNet in Figure 9 to highlight some of their differences.

The **Risk_scenario** in the field of the environment represents the potential threats to the ecosystem and some of its components. It also shows how the human (although responsible for most of these threats) takes measures to prevent some of them.

Interestingly, the **Risk_scenario** unveiled using data taken from an environment corpus and database shares some similarities, but also some differences with the one appearing in FrameNet. For

instance, the **Wagering** frame (that comprises LUs such as *bet* and *wager*) was completely irrelevant for the environment. Conversely, a **Preserve_in_original_state** was defined for the environment data for terms such as Eng. *conserve*, and Fr. *conservation*, *préserver*.

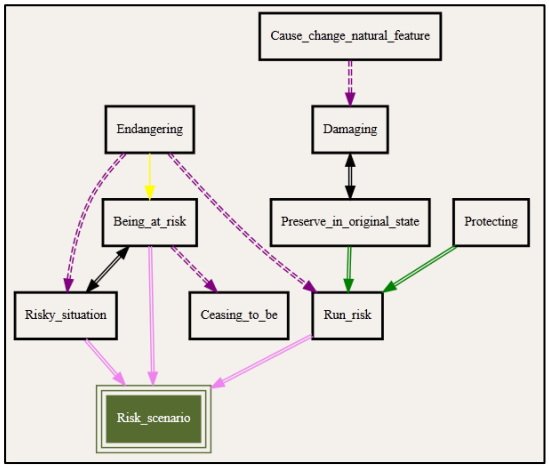


Figure 8. **Risk_scenario** in the environment

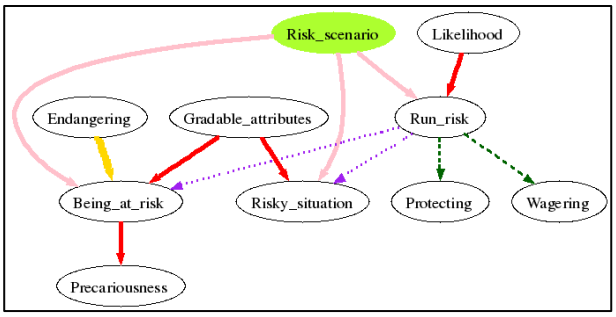


Figure 9. **Risk_scenario** in FrameNet (2014)

5.2 Managing_waste

The **Managing_waste** scenario was also defined based on the terms related to residual waste management. This scenario shows the different processes involved in managing waste and the order in which they are performed: first waste is collected, then it is separated; afterwards, it can be recovered or discarded. If waste is removed, it can then undergo incineration or landfilling. On the other hand, if waste is recovered, it is either recycled, composed or processed.

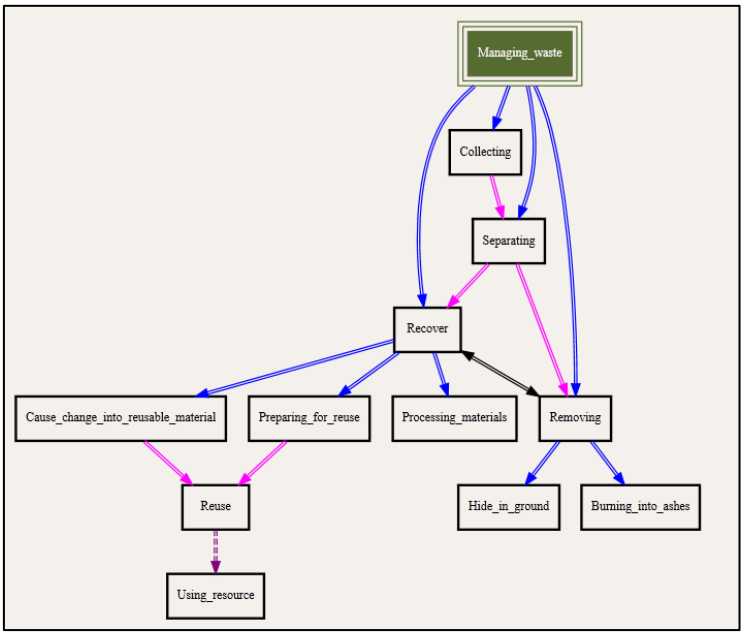


Figure 10. **Managing_waste** in the environment

6 Conclusion

In this paper, we presented a methodology to discover alternative conceptual structures for terminology. They complement structures often used to represent entity concepts (i.e. domain ontologies) and are well suited to account for terms denoting processes, events, and properties.

The methodology, based on principles borrowed from Frame Semantics and its implementation in FrameNet, was applied to English and French terms that are related to the field of the environment. It allowed us to unveil frames that are similar to those recorded in FrameNet, but also new ones that might be specific to the specialized field we chose to describe. It also allows us to represent small conceptual scenarios.

Acknowledgements

This work was supported by the Social Sciences and Humanities Research Council (SSHRC) of Canada. We would like to thank Carlos Subirats for his help during the first stages of this analysis and the three anonymous reviewers for their useful comments on a previous version of this paper.

References

Baker, C. & J. Hung. 2010. *Release 1.5 of the FrameNet data*. International Computer Science Institute. Berkeley.

Boas, H. C. (ed.) 2009. *Multilingual FrameNets in Computational Lexicography*. The Hague: Mouton.

Condamines, A. 1993. Un exemple d'utilisation de connaissances de sémantique lexicale : acquisition semi-automatique d'un vocabulaire de spécialité. *Cahiers de lexicologie* 62: 25-65.

Cruse, A. 2011. *Meaning in Language*. Oxford: Oxford University Press.

Dolbey, A., M. Ellsworh and J. Scheffczyk. 2006. BioFrameNet: A Domain-specific FrameNet Extension with Links to Biomedical Ontologies. *KR-MED 2006 Biomedical Ontology in Action*, Baltimore, Maryland.

Faber, P. et al. 2006. Process-oriented terminology management in the domain of Coastal Engineering. *Terminology* 12(2): 189-213.

Faber, P. (ed.). 2012. *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin/Boston: De Gruyter.

Fillmore, C. 1982. Frame Semantics. *Linguistics in the Morning Calm*, Seoul: Hanshin Publishing Co.: 111-137.

Fillmore, C.J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6(2): 222-254.

Fillmore, C.J. & C. Baker. 2010. A frames approach to semantic analysis. *The Oxford Handbook of Linguistic Analysis*, Bernd Heine and Haiko Narrog (eds): 313-339. Oxford: OUP.

Fillmore, C.J.. C.R. Johnson & M. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography* 16(3): 235-250.

FrameNet [https://framenet.icsi.berkeley.edu/fndrupal/]. Accessed 16 May 2014.

L'Homme, M.C. 1998. Le statut du verbe en langue de spécialité et sa description lexicographique. *Cahiers de lexicologie* 73(2): 61-84

L'Homme, M.C. 2012. Adding syntactico-semantic information to specialized dictionaries: an application of the FrameNet methodology. Gouws, R. et al. (eds.). *Lexicographica* 28: 233-252.

L'Homme, M.C. 2014. Terminologies and taxonomies, Taylor, J. (ed.). *Handbook of the Word* Oxford: Oxford University Press.

L'Homme, M.C., B. Robichaud & C. Subirats. 2014. Discovering frames in specialized domains, In *Language Resources and Evaluation*. LREC 2014, Reykjavik, Iceland.

Lerat, P. 2002. Qu'est-ce que le verbe spécialisé ? Le cas du droit. *Cahiers de Lexicologie* 80: 201-211.

Lorente, M. 2002. Verbos y discurso especializado. *Estudios de lingüística española* (ELiEs) 16.

Mel'cuk, I., A. Clas & A. Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*, Bruxelles: Duculot.

Montiel-Ponsoda, E., G. Aguado de Cea, A. Gómez-Pérez, & W. Peters. 2010. Enriching Ontologies with Multilingual Information. *Journal of Natural Language Engineering* 17: 283-309.

Pimentel, J. 2013. Methodological bases for assigning terminological equivalents. A contribution. *Terminology* 19(2): 237-257.

Ruppenhofer, J., M. Ellsworth, M. Petruck, C. Johnson & J. Scheffczyk. 2010. FrameNet II: *Extended Theory and Practice.* 2010. [http://framenet.icsi.berkeley.edu]. Accessed 16 May 2014.

Schmidt, T. 2009. The Kicktionary – A Multilingual Lexical Resources of Football Language. Boas, H.C. (ed). *Multilingual FrameNets in Computational Lexicography. Methods and Applications*, Berlin/NewYork: Mouton de Gruyter: 101-134.

Wandji, O., N. Grabar & M.C. L'Homme. 2013. Discovery of semantic frames for a contrastive study of verbs in medical corpora. *Terminology and Artificial Intelligence*, TIA 2013, Paris.