TERMINOLOGY and LEXICOGRAPHY RESEARCH and PRACTICE 20

Lexical Semantics for Terminology

An introduction

Marie-Claude L'Homme

John Benjamins Publishing Company

Lexical Semantics for Terminology

Terminology and Lexicography Research and Practice (TLRP)

ISSN 1388-8455

Terminology and Lexicography Research and Practice aims to provide in-depth studies and background information pertaining to Lexicography and Terminology. General works include philosophical, historical, theoretical, computational and cognitive approaches. Other works focus on structures for purpose- and domain-specific compilation (LSP), dictionary design, and training. The series includes monographs, state-of-the-art volumes and course books in the English language.

For an overview of all books published in this series, please see *benjamins.com/catalog/tlrp*

Editors

Marie-Claude L' Homme
University of Montreal

Kyo Kageura
University of Tokyo

Volume 20

Lexical Semantics for Terminology: An introduction
by Marie-Claude L'Homme

Lexical Semantics for Terminology

An introduction

Marie-Claude L'Homme
University of Montreal

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/tlrp.20

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2019040102 (PRINT) / 2019040103 (E-BOOK)

ISBN 978 90 272 0467 7 (HB)
ISBN 978 90 272 0468 4 (PB)
ISBN 978 90 272 6178 6 (E-BOOK)

John Benjamins Publishing Company · www.benjamins.com

To two mentors, *Jean-Claude Boulanger and Juan Carlos Sager*

Table of contents

List of figures

List of tables

List of abbreviations

Adj	Adjective
AP	Adjective phrase
Ar	Arabic
BNC	British National Corpus
Cs	Czech
ECD	Explanatory Combinatorial Dictionary
ECL	Explanatory Combinatorial Lexicology
En	English
Es	Spanish
Eu	Basque
FE	Frame element
Fr	French
FS	Frame Semantics
GTT	General Theory of Terminology
Ja	Japanese
LF	Lexical function
LU	Lexical unit
MTT	Meaning-Text Theory
N	Noun
n. f.	Noun, femimine (for French terms)
n. m.	Noun, masculine (for French terms)
NP	Noun phrase
PP	Prepositional phrase
Prep	Preposition
V	Verb
Vi	Intransitive verb
VP	Verb phrase
Vt	Transitive verb
Zh	Chinese

Typographical conventions

**sentence*	Unacceptable sentence
?sentence	Abnormal sentence
word or sentence in italics	Linguistic unit or expression
$word_1$, $word_2$	Lexical unit; lexical item with a specific meaning
$word_N$, $word_{Adj}$	Lexical item labeled with a part of speech; used when there is ambiguity
WORD in small capitals in example	Lexical unit or term being exemplified
word in italics and bold	Important notion that is explained (it also appears in the index)
'word between quotation marks'	Concept or meaning
~	In a collocation, the base or keyword
word ↔ ***word***	Lexical units sharing a semantic relation
word ⇔ ***word***	Interlinguistic equivalents
EDUCATION_TEACHING	Name of semantic frame
f(lexical unit) = linguistic unit	f: lexical function; (keyword); = value of a lexical function when applied to a lexical unit
X, Patient, Locative, Ingestible	Label used for arguments, circumstantials and frame elements
[The climate $_{Patient}$] WARMS	Argument or circumstantial and its linguistic realization in an annotated sentence
<is-a>	Relation between two concepts

Acknowledgments

I realized that understanding words was a passion of mine as soon as we were asked to identify synonyms and antonyms as a method to expand our vocabulary in primary school. Later, I was fortunate enough to meet other people who shared the same passion: professors (in terminology, computer science, lexicology or lexical semantics), colleagues and students with whom I shared many discussions on lexical units, meaning, terms, and ways to describe them in printed dictionaries or electronic resources. I am indebted to all these people and it would be impossible to name them all here. Nevertheless, I would like to express my acknowledgments to those who have more directly influenced the contents of this book:

- Jean-Claude Boulanger, Pierre Auger, Rostislav Kocourek and Juan C. Sager who made me realize that there is more to terminology than meets the eye.
- Many colleagues around the world who work in terminology and gave me opportunities to approach topics in the field from different perspectives, especially Béatrice Daille, Kyo Kageura, Patrick Leroyer, colleagues I met at the Universidad Politécnica de Madrid, Universitat Pompeu Fabra, Universidad de Granada, Université Toulouse-le-Mirail, Università degli studi di Verona, and, finally, authors of articles published in the journal *Terminology*.
- Current and former members of the Observatoire de linguistique Sens-Texte (OLST) for their willingness to give advice, spend time answering difficult questions, implement lexical descriptions, put up with my multiple requests and sometimes unrealistic ideas, especially, Patrick Drouin, François Lareau, Elizabeth Marshman, Igor Mel'čuk, Alain Polguère and Benoît Robichaud.
- Carlos Subirats who seems to never tire of discussing Frame Semantics and cognitive linguistics.
- Many students on my research team or in my classes whose questions often compel me to consider topics from an angle that had not previously occurred to me.

I would like to express my most heartfelt thanks to all those who contributed in one way or another to the development of the DiCoEnviro, the DicoInfo and the Framed DiCoEnviro. I also wish to thank two anonymous reviewers, Antonio San Martín and Victoria Stecum who carefully read a previous version of my book and pointed out different ways to improve its contents and/or layout.

CHAPTER 1

Why apply lexical semantics in terminology?

Terminology is the study of terms, and although terms are types of words, terminologists devised principles and methods to describe them that differ from the ones used in lexicology or lexical semantics. Terminology has crafted its own definitions for 'word', 'term' and 'meaning' and has traditionally held a perspective on linguistic units that contrasts with that of most linguists. The reason for this stems from the fact that terminology, on the one hand, and lexicology or lexical semantics, on the other, have vastly different objectives. These differences will become clearer in the next chapter.

Things are changing though and an increasing number of researchers (Condamines 1993; Lerat 2002a; Gaudin 2003; Aldestein and Cabré 2002; Faber and L'Homme 2014; among others) stress the usefulness of ***lexical semantics*** for terminology. In fact, lexical semantics and terminology have much in common, since both disciplines aim to answer questions about the nature of words, word content (i.e. the nature of meaning), the relationship between word content and our construal of reality, and relations between words and word meanings.

This book aims to show how the principles borrowed from lexico-semantic frameworks and methodologies derived from them can help us understand terms. It explains how lexical analysis can be integrated in most steps of terminology work and how it complements perspectives entirely focused on knowledge. It seeks to provide answers to questions such as those that are listed below.

- In a specialized text, which units or expressions correspond to terms? In a text on climate change, are *climate*, *climate change*, *change*, *emit*, *greenhouse gas*, and *warming* all relevant terms? How can terms be distinguished from other units?
  - Chapter 4 addresses this question.
- Which units should be included in specialized dictionaries or other kinds of terminological resources? In a dictionary pertaining to climate change, *climate* and *carbon dioxide* would normally be included. Should *warming* and *reduction* be considered as well? Should the verbs *change* and *fluctuate* and

the adjectives *extreme* and *renewable* also be added to word lists of specialized resources?
  - This issue is discussed in Chapters 4 and 5. Chapter 6 further explains that terms belong to different categories.
- When units have multiple meanings, are all meanings relevant for specialized resources? *Environment* can be defined as: 1. 'the circumstances, objects, or conditions by which one is surrounded' (Merriam-Webster Online Dictionary 2018); 2. 'the complex of physical, chemical, and biotic factors (such as climate, soil, and living things) that act upon an organism or an ecological community and ultimately determine its form and survival (Merriam-Webster Online Dictionary 2018)'; 3. 'the specific configuration of a computer system'. In an environmental dictionary, only the second meaning would be of interest. Furthermore, different meanings can pertain to the same subject field. *Address* in computer science can refer to: 1. 'the unique identifier of a user'; or 2. 'an identifier of a location on a storage medium'. These two meanings must be described in a computer science dictionary. How can we distinguish meanings and ensure that we are dealing with separate meanings?
  - These questions are addressed in Chapter 5.
- When considered from the perspective of a special subject field, does the meaning of a word differ from its most common meaning? Using the *address* example above, the meanings it conveys in computer science are not entirely compatible with the meaning laypeople would initially think of, which can be roughly paraphrased as 'the name of a place where a person lives or an organization is located'. Similarly, does *write* in *write a program in Java* differ from the *write* in *write a letter in English*? Should only meanings that are not encompassed by general language be considered in specialized dictionaries?
  - This will be discussed in Chapter 5.
- How can we tell if words are related? Are *global warming* and *climate change* related? Of course they are somehow, but can they be used interchangeably? What is the opposite of *green* in the environment? Is it *polluting* or *fossil*? *Pollute*, *pollution*, *polluted*, *polluting*, and *polluter* are related but in different ways; how can we describe these relations? How can we account for the fact that the verb *pollute* and the noun *polluter* are related the same way the verb *program* and the noun *programmer* are?
  - These questions and many others are the topics of Chapters 7 and 8.
- How can we explain the fact that words do not combine randomly? To express the idea of 'creating a program' in computer science, which verb would be adequate? *Write* is an option. However, with other nouns, the same meaning would be expressed with different verbs: e.g. *create* or *define a password*, *create a file*, *develop an application, type a message*.

  - This matter will be addressed in Chapters 7 and 8.
- How are the same meanings expressed in different languages? It is possible to establish straightforward equivalence relationships between sets of terms (e.g. En: *computer* ⇔ Fr: *ordinateur* ⇔ Es: *ordenador* ⇔ Ja: コンピューター; En; *environment* ⇔ Fr: *environnement* ⇔ Es: *medio ambiente* ⇔ Ja: 環境). However, some equivalents have different structures in their respective languages: e.g. the verb *to bookmark* translates into French as a collocation, namely *mettre en signet*.
  - This issue is discussed in Chapter 9.
- Are meanings always expressed similarly in different languages? While this is often an assumption for specialized terms, there are counter-examples: the English term *farming* may be rendered in French by *élevage* or *culture* according to the objects being farmed. *Élevage* usually applies to animals, while *culture* concerns plants.
  - The matter is also dealt with in Chapter 9.

This book shows how the answers to these fundamental questions can be found in the principles, methods and tools developed by lexical semantics. Before answering these questions, Chapter 2 examines the paradigm to which terminologists or other experts involved in terminological activities have often referred. Since it remains a very important strand in terminology circles, most subsequent chapters refer to this paradigm and point out some of the questions it raises. Chapter 3 will show how lexical semantics can help terminologists find answers to some of these questions. The following chapters highlight various ways to differentiate lexical semantics from this more traditional approach.

Even if lexical semantics is useful for terminology, it is more directly relevant to some of its applications. Any task that deals with terms in running text – with or without automated methods – needs to implement some degree of lexico-semantic analysis. Such applications include – but are not limited to – specialized dictionary compilation, specialized translation, semi-automated or automated processing of specialized texts, and document indexing. Other applications, such as knowledge modeling, standardization, language planning, may be concerned with the lexico-semantic properties of terms, but only during preliminary stages, i.e. when analyzing the meaning of specialized lexical items. These applications distance themselves from purely linguistic considerations when making decisions on how knowledge must be represented or which terms are more suitable than others in a given communicative setting.

This book is intended for readers who do not necessarily have a background in terminology. But it can also be of interest for readers who already have a background in terminology but not in lexical semantics. It should complement other

textbooks that do not focus on lexical semantics *per se* (Cabré 1999; Dubuc 2002; L'Homme 2004b; Pavel and Nolet 2001; Sager 1990; for instance). Lexicographers and linguists or any other person interested in words can also learn from the specific issues raised by the description of specialized lexica.

This book is intended for readers who wish to acquire a broad perspective on the possible connections between lexical semantics and terminology. Bibliographical pointers are given at the end of each chapter to works that provide more detailed analyses on the different studied topics.

Illustrative examples taken mainly from the fields of computing and the Internet, the environment, and cycling are used throughout the book, since these are the fields with which I am most familiar. Examples regarding computing and the Internet, on the one hand, and in environment, on the other, are taken from two online resources, namely the DiCoInfo (2018) and the DiCoEnviro (2018) that my research group has been compiling during the past years. The book also makes reference to terms that belong to other fields based on work in terminology literature. However, the principles and methods presented should be applicable to many other fields of knowledge.

CHAPTER 2

Terminology

This chapter briefly describes terminology, its basic principles and the knowledge (or conceptual) paradigm in which terms are still often considered.[1]

Firstly, ***terminology***[2] studies ***terms*** which can be defined as linguistic expressions that designate ***items of knowledge*** within ***special subject fields***. There are several important keywords in the previous sentence that require further explanations.

– Terminology considers linguistic units from the point of view of ***special subject fields*** or ***domains*** (such as chemistry, computer science, cycling, environment or law). The notion of 'special subject field' is defined rather loosely in terminology and often corresponds to a scientific discipline (e.g. biology, economics, or linguistics) or a technology (e.g. Internet, imaging, or auto mechanics). Economic sectors can also be considered special subject fields in terminology (e.g. recycling, steel machining) as well as schools of thought (e.g. cognitive or structural linguistics). Even sets of objects have been defined as subject fields. Wüster, the founder of terminology, compiled a dictionary entitled *The Machine Tool Dictionary* (1968) dealing with types of machine tools, their parts, etc. The delimitation and classification of subject fields are usually linked to the objectives of those who carry out these delimitations and/or classifications. The subject field (or *domain*) sets boundaries within which terminologists analyze linguistic expressions: only those units that can be connected to a domain are considered relevant.

---

1. Readers who are familiar with the traditional principles of the field can skip the next two sections or browse through them. Section 2.3, however, lists some questions raised by this paradigm, especially when considering terms and terminological analysis from a linguistic point of view.

2. This section defines 'terminology' as a field that studies terms and considers them from the point of view of applications. It should be noted that *terminology* is polysemous and can also be used to designate the set of terms of a given domain. For instance, *environmental terminology* refers to all the terms found in that field of knowledge.

- It is assumed that subject fields – or more specifically the knowledge they encompass – have a ***structure***. Terminologists take this structure into account when describing terms.
- The structure of a subject field first consists of ***items of knowledge***, called ***concepts***.[3] ***Concepts*** are defined according to the place they hold in a conceptual structure.
- ***Conceptual structures*** consist of ***concepts*** and ***relations*** that concepts share. Common relations considered in terminology are taxonomic, often expressed with <is-a> (e.g. the printer <is-a> device), and partitive, expressed with <part-of> (e.g. a planet is a <part-of> a solar system).
- Concepts are labeled with linguistic expressions called ***designations***. The association of a concept with a designation produces a ***term***.[4] For example, the concept 'device used in conjunction with a computer to reproduce data on paper or another medium' is labeled with *printer*; 'astronomical object revolving around a star' is designated by *planet*.

Secondly, terminology is deeply rooted in ***applications***, such as specialized dictionary compilation, specialized translation, document indexing and/or classification, knowledge modeling, language planning, and standardization. This means that terminology (its theoretical and methodological principles) aims first and foremost to provide answers to the questions raised by these applications. This firm grounding in applications distinguishes terminology from other branches of linguistics, such as lexicology or lexical semantics that can study words regardless of possible applications. Thus, a lexicologist can analyze words or selected subsets of words without ever thinking of compiling a dictionary. A terminologist, on the other hand, considers terminological data within at least one of the applications mentioned above.

We can see that the applications in which terminology principles are used are diverse. Knowledge modeling seeks to represent knowledge in a formal manner and aims to label items in knowledge structures unambiguously. The purpose of document indexing is to identify units (i.e. terms or ***descriptors***) that are most representative of the contents of a book or article, standardize these units, and store them in a repository where users can access them. Hence, there

3. Here, I will not elaborate further on the complicated notion of 'concept'. Different views on 'concept' and 'meaning' are presented in Section 3.1.1 and in Chapter 5.

4. In fact, *designation* is only used when the context requires it. Often, authors use *term* instead of *designation* when referring to the linguistic expression used to label a concept. This shortcut will also be taken in this book.

are different ways to carry out terminology work, and this inevitably has affects on the way the notions of terminology are considered. The next section presents the most widespread paradigm in which terminology is situated. This paradigm has been challenged more than once and other perspectives offer alternative ways to view terms.

2.1 The knowledge paradigm[5]

The creation of terminology as a discipline is attributed to ***Wüster*** (1979). However, the principles presented in this section result from Wüster's ideas and changes that occurred in the years that followed his first proposal (in the thesis he defended in 1931) as well as reinterpretations of his writings (Humbley 2007; Campo 2013). These principles fall under the scope of the ***General Theory of Terminology (GTT***), now often called ***traditional terminology***. In addition, previous work in other fields which involved organizing scientific knowledge in structures and developing methods of standardization had been carried out before Wüster. The main contribution of Wüster and his followers was to formulate sound theoretical principles to frame this work.

Wüster designed a theory in order to meet objectives chiefly targeted at providing means to ensure unambiguous communication. Two main assumptions are associated with the GTT:

- Knowledge has a ***structure***. Hence, linguistic units used to convey this knowledge reflect that structure.
- ***Specialized communication*** needs to rely on ***unambiguous linguistic units***.

Regarding the first assumption (= knowledge has a ***structure***), it parallels the many efforts made in scientific disciplines to organize and represent knowledge consistently and explicitly. The most famous example of this is the ***taxonomy*** of animal species in zoology (i.e. the principles of which have been agreed upon among experts).

According to these principles, species are organized based on their natural relationships (an example is shown in Figure 2.1). A complex system of levels, such as class, order, family, and genus, accounts for all known species and tells us how similar or different one species is from the next. There are more examples of

5. The material presented in this section has been edited and adapted from that printed in L'Homme (2015a).

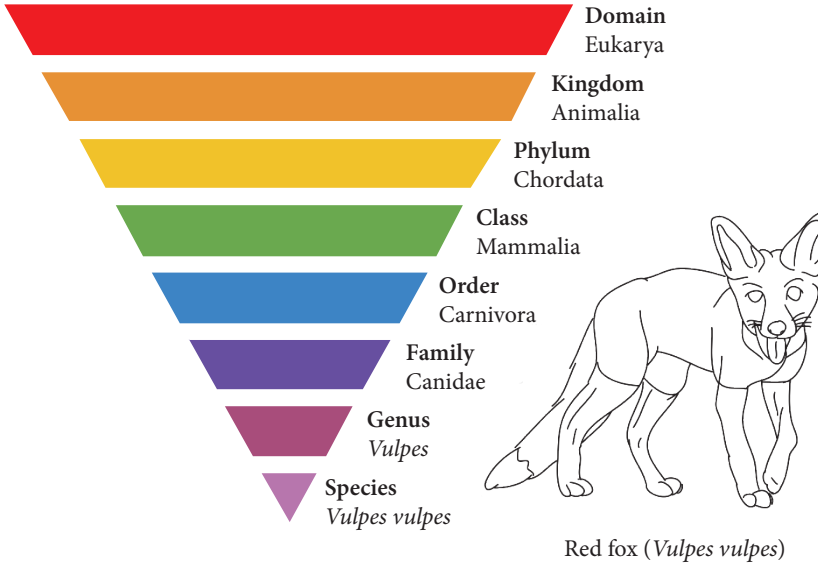


Figure 2.1 Classification of *Vulpes vulpes* (commonly known as *red fox*) (Wikimedia Common 2018 (https://commons.wikimedia.org/wiki/File:Taxonomic_Rank_Graph.ai))

knowledge classification in other domains. For instance, the periodic table classifies chemical elements according to their atomic mass.

The second assumption (= ***specialized communication*** needs to rely on ***unambiguous linguistic units***), is also in line with previous efforts made by scientific communities. Coming back to the zoology example, experts – more specifically the International Commission on Zoological Nomenclature – defined a series of rules to establish standard names for animals. These rules can be found in the *International Code of Zoological Nomenclature* (2000). The Code determines what names are valid for a ***taxon*** (an item in the taxonomy) at the family, genus, and species levels. Naming conventions apply the principle of homonymy (the name of each taxon must be unique), the principle of binominal nomenclature (a species name is a combination of a generic name and a specific name), and the principle of coordination (the Code 2000). Other guidelines are meant to deal with competing designations for the same taxon.

Other scientific disciplines also use names that label items of knowledge unambiguously. In chemistry, symbols reflect the composition of elements and compounds (H_2O for *water; Ti* for *titanium).* Also, for reasons of international comprehensibility, medical terminology favors Greek and Latin morphemes with specific meanings in the creation of terms (e.g. *-itis* used to designate an inflammation as in *hepatitis* and

meningitis). For obvious reasons (namely, the time required to define and agree on criteria and rules), a small number of domains have very clear guidelines for classifying and naming entities and the above examples pertaining to zoology, chemistry and, to a lesser extent, medicine, can be considered success stories in that respect.

The General Theory of Terminology contributed to these matters by formulating theoretical and methodological principles for classifying components of knowledge and conventionalizing names. Some of these principles are listed below.

As was said above, in terminology, items of knowledge are called ***concepts***. These are assumed to be generalizations of extralinguistic entities, i.e. the items of knowledge manipulated in science and technology (concrete objects such as 'computer', 'wind turbine', representational entities such as 'data', animate entities such as 'environmentalist', activities such as 'decontamination').

Traditional terminology theory postulates that concepts are the starting point of the analysis of terms. Hence, terminologists should start by identifying concepts and then find or assign designations to them. Figure 2.2 illustrates the application of this ***onomasiological approach***. This entails that concepts can – at least theoretically – be regarded as independent entities that can be delineated regardless of the linguistic expressions that serve to label them.

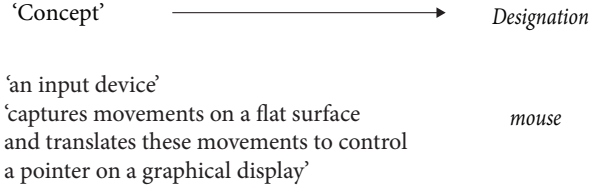


Figure 2.2 Onomasiological approach

Concepts are delineated by methods of classification similar to those used in zoology. For example, when defining the concept 'mouse' in the field of computing, terminologists must distinguish it from other neighboring concepts. The result of this classification is called a ***conceptual structure***. Figure 2.3 shows that 'mouse' falls under a generic concept, i.e. 'input device', that it shares a certain number of characteristics with 'keyboard' since it has the same superordinate concept. It also differs from 'keyboard' since the two concepts each have characteristics that set them apart.

More can be said about the position of 'mouse' in this conceptual structure. 'Mouse' subsumes two specific concepts, namely 'optical mouse' and 'wireless mouse'; they both possess all the characteristics attributed to the concept 'mouse' but have one or two additional characteristics. Finally, 'mouse' shares a part-whole

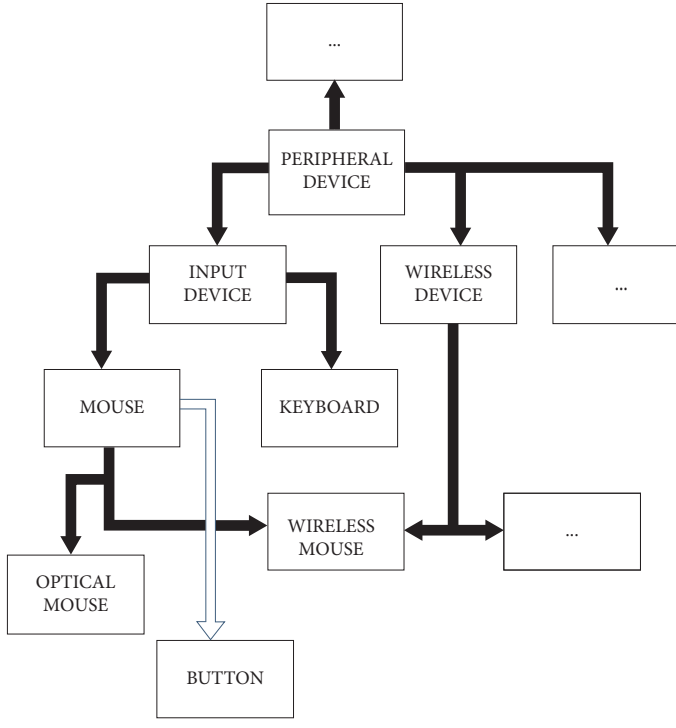


Figure 2.3 'Mouse' and related concepts in a miniature conceptual structure

relation with 'button'. This simple conceptual structure illustrates the two main types of relations studied in terminology (taxonomic and partitive). However, many more relations exist between concepts. When describing terms in dictionaries, i.e. when defining them, terminologists should make these relations explicit.

> Mouse: 'Input device' operated by moving it on a flat surface and used to control the cursor and select icons and files on a computer screen.
> Keyboard:'Input device' that consists of a series of keys used to enter data into a computer.
> Wireless mouse:'Mouse' with no cord that transmits data to a computer via infrared radiation.
> And so on.

Of course, this approach assumes that concepts can be defined precisely, at least at a certain point in time. If concepts change in the way they are delimited, their definition will inevitably change and this might affect their position in the conceptual structure.

Another important principle in traditional terminology is that of ***univocity***. According to this principle, a concept should have a single designation;

conversely, a designation should be used to label a single concept. This means that ***synonymy*** and ***polysemy*** should be avoided. Of course, this ideal situation is often challenged and terminologists must create rules similar to those used in zoology or adapt them to specific situations. For instance, the concept 'greenhouse gas' defined loosely as 'a gas that absorbs and emits radiation from the Earth's surface' can be labeled *greenhouse gas*, *GHS*, and *greenhouse effect gas*. Regarding polysemy, *address* in computing can designate: 1. 'the unique identifier of a user'; 2. 'an identifier of a location on a storage medium'. In terminology, polysemy is less prevalent than in general language since terminologists concentrate on meanings within specialized domains; synonyms, on the other hand, are quite common.

Although rules such as those applied in zoology exist but in a few subject fields, some regularity can be observed in the terminologies of many domains (Sager 1990). This is due in part to a willingness to use transparent (i.e. compositional) terms. ***Transparent terms*** facilitate the understanding of new concepts and probably contribute to the success of knowledge transfer. The forms of many terms reflect the concepts they denote by stating some of their characteristics, such as a function, a component, or an origin. For instance, *input* in *input device* means that the device is used for entering data into the computer. *Optical* in *optical mouse* designates the way the peripheral captures movements (by means of an optical sensor). Morphemes can have the same function: *bio-* in *biodiesel*, *bioethanol*, *biofuel, biogas*, and others indicates that the term denotes something that is made of organic matter. Terms can also reflect the place of concepts in a hierarchical structure as shown in the examples below in which terms designate more and more specific concepts:

file → installation file → program installation file
energy → wind energy → offshore wind energy

Designations can make a difference between two or more concepts explicit (e.g. *sustainable development <-> unsustainable development; afforestation <-> deforestation <-> reforestation*). Finally, another common mechanism to create transparent terms is to resort to metaphors (e.g. *anchor*, *memory* or *mouse* in computing; *greenhouse gas* in the environment).

2.2 Storing and accessing concepts and terms

Different repositories can be used to store terms, preferred terminological choices, and represent conceptual relations. The results of the analysis carried out by terminologists are recorded in ***terminological databases***, ***term banks*** or

specialized dictionaries. Many are structured in a way that differs from general language dictionaries to comply with the onomasiology and univocity principles explained in Section 2.1.

A ***term record*** is designed to account for the description of a concept and a single one. All designations that express this concept should appear on the same record. Figure 2.4 shows how this is done in the TERMIUM Plus® term bank for the concept 'global warming'. In French, various designations can label the concept defined as 'The rise of temperature of the atmosphere as a consequence of the greenhouse effect': i.e. *réchauffement climatique*, *réchauffement de la planète*, *réchauffement planétaire*, etc. Often, one of these designations is preferred and the other ones are listed as synonyms, variants, or deprecated forms. Although the TERMIUM Plus® record does not indicate explicitly a preferred term (Figure 2.4), *réchauffement climatique* is placed first and some expressions are labeled as expressions to be avoided. *Réchauffement du globe* is labeled as obsolete; *réchauffement global*, as an anglicism.

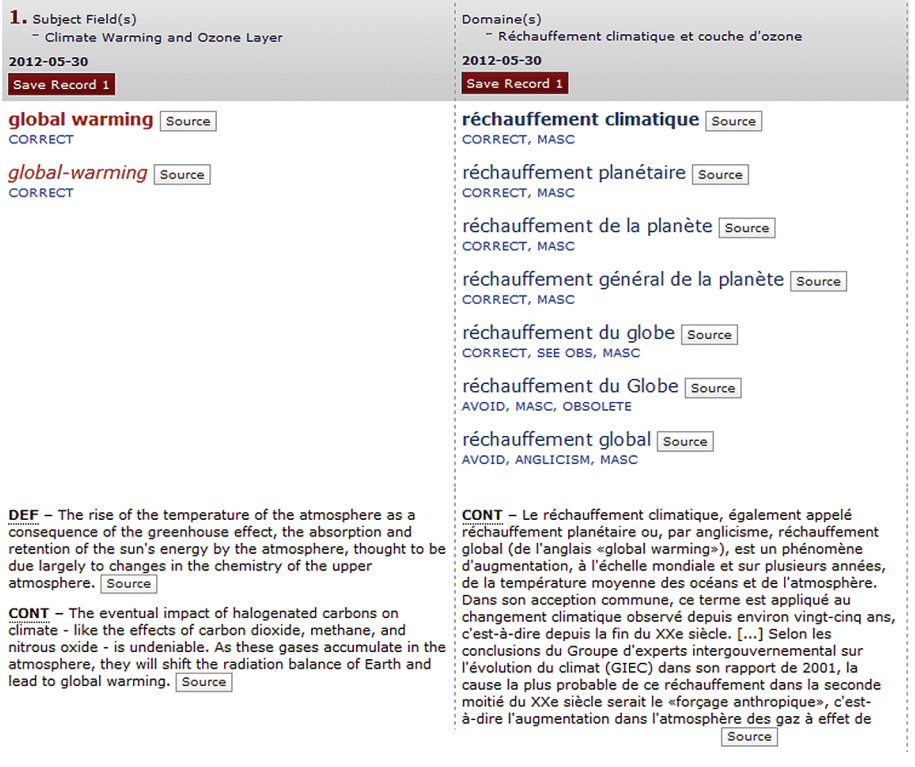


Figure 2.4 First part of the term record 'Global warming' in TERMIUM Plus® (2015)

Even if the different data categories (part of speech, usage note, definition, context, etc.) are organized in a consistent manner in term banks, relations between terms or more specifically between the concepts for which they stand are seldom properly formalized. They are implicit rather than explicit and they must be inferred by users.

Other repositories offer better ways to represent relations between concepts. One example is the ***thesaurus*** that implements hierarchical relations. The thesaurus also takes into account other links under the general heading 'associative relations'. Figure 2.5 is an entry extracted from a thesaurus on the environment, called GEMET (2015). In this thesaurus, the preferred equivalents in many different languages are also listed. Hierarchical relations appear under the labels 'broader terms' and 'narrower terms'.

Habitat

Definition:
(1) The locality in which a plant or animal naturally grows or lives. It can be either the geographical area over which it extends, or the particular station in which a specimen is found. (2) A physical portion of the environment that is inhabited by an organism or population of organisms. A habitat is characterized by a relative uniformity of the physical environment and fairly close interaction of all the biological species involved. In terms of region, a habitat may comprise a desert, a tropical forest, a prairie field, the Arctic Tundra or the Arctic Ocean. (Source: WRIGHT / GILP)

broader terms
- synecology

narrower terms
- animal habitat
- wildlife habitat

Scope note:

scope note is not available

Groups:
BIOSPHERE (organisms, ecosystems)

Themes:
Biology
natural areas, landscape, ecosystems

Other relations:
Wikipedia article
Habitat
Has close match
UMTHES: Habitat

Arabic:	موطن
Basque:	Habitat
Bulgarian:	Ареал, местообитание
Catalan:	Hàbitat
Chinese:	栖息地
Croatian:	stanište
Czech:	stanoviště
Danish:	Habitat
Dutch:	Habitat
English (US):	Habitat
Estonian:	elupaik, kasvukoht
Finnish:	kasvuympäristö, elinpaikka
((…))	

Figure 2.5 Entry 'habitat' in GEMET (2015)

Conceptual relations can also be encoded in even more sophisticated formal structures called ***ontologies*** (Staab & Studer 2009).[6] A widely accepted definition of 'ontology' is worded by Studer et al. (1998): "An ontology is a formal, explicit specification of a shared conceptualization." In ontologies, knowledge items are defined according to a set of properties and linked via explicit relations. The ***representation*** of knowledge in ontologies (the ***conceptualization***) is said to be ***consensual***. It must account for criteria on which a community of experts agrees. In a way, ontologies represent the ideal structures in which the principles of the GTT can be implemented. Paradoxically, research on ontologies is an area that emerged and is still often carried out outside the realm of terminology *per se*.

Ontologies can account for various kinds of relations between concepts (taxonomic, but also many others such as partitive, cause-effect). Figure 2.6 shows how

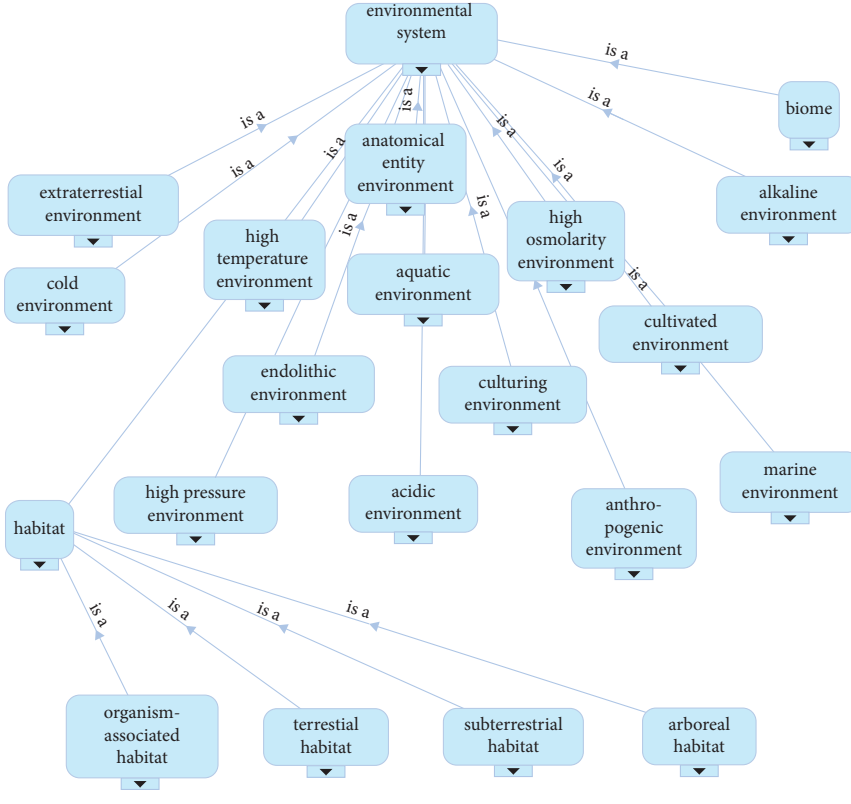


Figure 2.6 'Habitat' and related concepts in ENVO (2015): Graphical representation

6. A less formal version of ontologies can be found in ***terminological knowledge bases*** (Meyer et al. 1992).

'habitat' is represented in an ontology dedicated to the environment called *Envo*. The figure provides a view on taxonomic relations. Of course, more information can (and is) provided on each concept (a definition, variants, etc.). Since concepts and relations are represented formally, different kinds of operations can be carried out on them.

2.3 The knowledge paradigm from a linguistic point of view

Although the General Theory of Terminology (GTT) formulated principles that are compatible with the needs it was designed to fulfill, i.e. produce unambiguous naming conventions based on a consensual representation of knowledge, these principles impose a certain perspective on terms. Furthermore, they have linguistic consequences that are examined in this section.

In the GTT, terms are assumed to be linguistic labels that can be superimposed on knowledge configurations. Subject field experts or terminologists build consensual or *ad hoc* conceptual structures, such as those that appear in Figures 2.1, 2.3 and 2.6, and do this prior to reflecting on the linguistic forms used to express this knowledge.[7] By doing so, scientists or terminologists sometimes need to create categories (and labels for them) that would not have been spontaneously proposed if only natural distinctions had been at play. For instance, the 'order' and 'family' categories in Figure 2.1 or the 'wireless peripheral' category in Figure 2.3 are the result of classification efforts in zoology and computing. Similarly, 'anatomical entity environment' and 'high osmolarity environment' shown in Figure 2.6 were added to better organize knowledge in the environment and are certainly not a reflection of the use of linguistic expressions in running text.

The relevance of a linguistic expression to act as a term is assessed according to its capacity to label a concept and not according to its linguistic status. Hence, *Vulpes vulpes* is considered a relevant term since it labels a well-defined group of animals in the field of zoology. Although it is nearly never used in everyday language, used by experts in very specific situations and was created artificially, it is still a term. Similarly, *climate*, *greenhouse gas*, and *endangered species* are considered to be terms in the field of the environment even though they correspond to

7. The methods for delineating concepts and gradually unveiling conceptual structures presented above might be applied by experts. However, they can seldom be used systematically by terminologists since they are not experts in the fields they are asked to explore. Terminologists are not in a position to precisely delineate concepts without previously relying on the content of texts. They usually apply an iterative analysis that consists in reading texts and acquiring knowledge, and then gradually build a conceptual structure.

different units from a linguistic point of view: *climate* is a single-word lexical unit, *greenhouse gas* is a non-compositional phrase (hence, a multiword lexical unit) and *climate change* is a compositional phrase.[8]

Terms are viewed as canonical labels isolated from text. In extreme cases such as zoology or chemistry where the naming conventions aim at yielding systematic, official, and universal names, the result is a form of semi-artificial linguistic apparatus that serves as a reference in specialized situations. Of course, most units handled by terminologists are not artificial and pertain to natural language. However, to be standardized, terms must be defined according to the way they denote a concept (the concept itself being well-defined within a field) and not according to the way they behave in the language at large.

A strong implicit assumption of the knowledge paradigm is that concepts and conceptual structures are language-independent. This even led taxonomists to suggest international designations for concept classes, such as *Vulpes vulpes*. The assumption is shared by the compilers of terminological resources: designations in multiple languages are listed on the same record (Figure 2.4) or in the same entry (Figure 2.5). Even if languages resort to very different linguistic units to label concepts (for instance, a single-word term in language A and a multiword term in language B), the structure of units is not a concern when establishing links between concepts and designations.

The vast majority of terms are nouns. This is a consequence of the focus of terminology on concepts (most of them being entities) and the way 'concept' is approached. Even in cases where activity concepts (linguistically expressed by nouns or verbs) or property concepts (prototypically expressed by adjectives) need to be taken into account, nouns are still preferred. This led to the assumption that specialized discourse is predominantly nominal in character.

Designations tend to be complex and often take the form of derivatives, compounds or multiword units. This is the result of the preference for transparency and clarity in scientific communication. This is best illustrated in Figure 2.6: only two concepts in this part of the ontology are labeled with single-word terms, i.e. *biome* and *habitat*.

In recent years, the perspective advocated by traditional terminology has been challenged from many different angles (Béjoint and Thoiron 2000; Bourigault and Slodzian 1999; Cabré 2003; Diki-Kidiri 2000, 2007; Faber 2012; Gaudin 2003;

8. This being said, a selection is usually made among linguistic units that are considered to be terms even if they convey specialized knowledge. For instance, *endangered species* will normally be defined as a term, but not *this species is endangered*, which would most likely be defined as a variant. More will be said about variants further on.

Kageura 2002, 2012; Sager 1990; Temmerman 2000). Most critics attack the implicit assumption held by the GTT that terms should be considered outside their linguistic, sociolinguistic or cultural context. Many also highlight the fact that terms are found in running text and behave like other linguistic units. As Sager pointed out a few decades ago:

> [The] origins [of terms] in texts of genuine natural language must, however, never be forgotten because terminologists use texts as their basic material and the dictionary tools they develop are intended to explicate natural language items or advise on the usage of terms in natural language contexts. (Sager 1990: 55)

The criticism also started to emerge at a time when changes were affecting work in terminology or other related fields such as lexicography. One important change was the introduction of methods and tools for compiling and processing corpora. The GTT does not provide all the answers to questions raised by these changes. Furthermore, like other fields, terminology witnessed the circulation of new ideas in linguistics, lexical semantics, and cognitive science. Some of them prove useful for supporting terminological analysis, but are incompatible to some extent with the GTT.

Of course, one of the objectives of the General Theory of Terminology is to label pieces of specialized knowledge unambiguously. Clarity is certainly important in scientific communication, but there is a long way to go to foster unambiguous communication if the starting point is the language used to express it. Knowledge cannot be communicated fully on the sole basis of conceptual structures and labels for its components. Only language has the expressive power to convey all the richness, complexity and subtleties of human knowledge.

Summary

Terminology, when it was founded, sought to make communication less ambiguous. To this day, its principles and methods have been used in several organizations especially those involved in standardization, knowledge modeling and translation. They are applied to compile term banks, thesauri and, to a certain extent, ontologies.

However, the assumptions made in traditional terminology raise many questions especially when they are considered from a linguistic perspective. The principles defined by the General Theory of Terminology and the methodology derived from them consider terms as labels for concepts, and as units emptied of most of their linguistic properties.

Most approaches to terminology now recognize that terms are full linguistic units and as such appear in linguistic environments in which they may undergo the same kinds of phenomena to which general lexical units are subjected, such as variation and ambiguity. Instead of attempting to neutralize these phenomena, new approaches reflect on different ways to handle them or further characterize their contribution to knowledge transfer. Communicating knowledge (specialized or non-specialized) exploits the full expressiveness of language; furthermore, knowledge cannot be fully captured in a conceptual structure or any knowledge structure for that matter. Hence, the challenge that now faces terminologists is to find a balance between an ideal for unambiguous communication and the fact that language changes and accommodates vagueness as well as ambiguity quite naturally.

Chapter 3 examines what two approaches – a knowledge-driven approach and a perspective guided by the lexical content of specialized texts – can tell us about terms. It will be argued that both are necessary to fully understand the relationship between linguistic units and specialized subject fields but each handles terms differently.

Further reading

To learn more about the principles of the General Theory of Terminology, refer to Wüster (1979), Felber (1984), Picht and Draskau (1985) and Rondeau (1984). Wüster designed an English-French dictionary based on the principles of the GTT, entitled the *Machine Tool Dictionary* (1968). To this day, it remains the most systematic application of the principles advocated by the scholar.

Cabré (1999), Sager (1990), Dubuc (2002), L'Homme (2004b), and Pavel and Nolet (2001) are introductions to the theoretical and methodological principles of terminology. It should be stressed however that these textbooks are not all firmly grounded in the GTT.

The principles of the GTT and new ideas that have emerged in terminology since the 1990s are summarized in Cabré (2003) and L'Homme (2015a). Campo (2013) is a thorough analysis of the work by Wüster and its reception in terminological literature.

For an introduction on thesaurus construction, see Aitchison et al. (2000). Staab and Studer (2009) contains many contributions on ontologies. For insights on the connections between ontologies and terminology, see Montiel and Aguado de Cea (2012), Roche (2012) and Durán-Muñoz and Bautista-Zambrana (2017).

CHAPTER 3

Lexical semantics for terminology

The previous chapter described some of the principles entrenched in a paradigm that we labeled the *knowledge paradigm* and also highlighted some of its limitations. We will now consider terminological data from a different perspective. We will start by comparing the kinds of questions that a standard terminological analysis (driven by domain knowledge) and lexical semantics (driven by lexical units and the meanings they convey) try to answer while using a concrete example. Lexical semantics principles can guide terminology work in the sense that it handles some linguistic properties of terms that other terminological frameworks driven by knowledge overlook.

3.1 A basic illustrative example

A simple example will be used to show how terminologists consider terminological data in practice and the various questions they need to address in the process. In this example, a terminologist is asked to compile a computing dictionary. One of the steps to be taken consists in building a ***corpus*** with representative texts. The corpus may be processed with the help of computer applications, such as term extractors and/or concordancers, in order to identify relevant terms and find information about them. At one point or another during the process, the terminologist will inevitably run into the unit *program*.

Terminologists interested in domain knowledge will probably determine that this term corresponds to a concept loosely defined as 'a set of instructions written in a computer language designed to perform a task or several related tasks automatically (i.e. by means of a computer)'. Once delineated, this concept becomes the focal point of the analysis.

Next, the analysis should lead terminologists to notice that this concept can be labeled with two different designations in English, i.e. *program* or *computer program* (in addition to the graphical variants *programme* and *computer programme*). Furthermore, the analysis should reveal that *program* can also refer to another concept in the field: 'a series of programs designed to carry out specific

tasks defined within a given application'. Again, this new concept can be labeled with another term, i.e. *software program*. Table 3.1 summarizes the results of this preliminary analysis. In both cases, a preferred term can be chosen, but this selection can be postponed to a later stage. Finally, if different languages are taken into account, it should not affect the way concepts are delineated. Of course, designations are language-specific, but the concepts themselves remain the same.

Table 3.1 Concepts designated by *program*

'Concept'	**Term(s) in English**	**Term(s) in French**
'A set of instructions written in a computer language designed to perform a task or several related tasks automatically (i.e. by means of a computer)'	*program* (1) *computer program* *programme* *computer programme*	*programme* *programme informatique*
'A series of programs designed to carry out specific tasks defined within a given application'	*program* (2) *software program* *programme* *software programme*	*logiciel* *application* *logiciel d'application*

Another important task consists in identifying the relations between 'program (1)' and other concepts. The concept 'program (1)' belongs to the larger class of 'software'. It includes more specific concepts, such as 'background program', 'backup program', 'bridge program', 'client program', and many more. Other concepts that refer to parts of a program can also be identified at this stage: 'instruction', 'loop', 'routine'. Taxonomic and partitive relations are commonly represented in terminology, but other conceptual relations can also be taken into consideration: a program is created with a specific means, i.e. a language; a program is designed to accomplish something, a 'task'. Ideally, important conceptual relations should be expressed in definitions.

Figure 3.1 depicts the process described above with *program* applied to a different concept, i.e. 'bird'. The objective of this kind of analysis is to delineate concepts precisely, situate them in a conceptual structure and label these concepts with designations. Designations often correspond to linguistic expressions found in specialized texts, but they do not necessarily have to since a new label could be chosen as a better way to represent a concept.

Let us now analyze *program* using a different starting point, a set of occurrences of *program* found in a corpus of texts on computing. One of the first things terminologists should do is discriminate the attestations of the verb

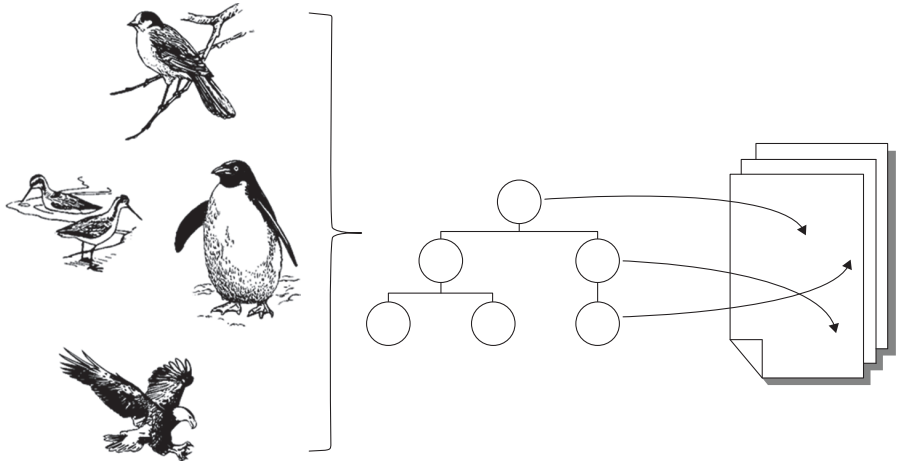


Figure 3.1 A schematic representation of a knowledge-driven approach to terminological data

program and the noun *program*. Both parts of speech appear in Table 3.2: lines 1, 2, and 6 show occurrences of the noun, while lines 3, 4, and 5 contain occurrences of the verb.

Table 3.2 Concordances with *program*

1	An FTP (file transfer protocol)	program	like Fetch on your hard drive helps locz
2	components available. Writing a	program	that works perfectly with every combinati
3	ot. First, students know how to	program	but do not know how to use the tools of
4	de, you'll need to learn how to	program	or rely on someone who can. Otherwise, y
5	cessors are generally easier to	program	than their fixed point cousins, but usual
6	lly, you can find the Uninstall	program	by clicking on the Start button at the lo

A larger set of concordances extracted from a corpus on computing should reveal that different meanings for each part of speech need to be distinguished.

> *program*$_1$, n.: 'A set of instructions written in a specific language …'
> *program*$_2$, n.: 'A series of programs designed to carry out specific tasks defined …'
> *program*$_3$, vt: 'To write a series of instructions using a computer language in order to have a computer carry out a task or a series of tasks automatically'.
> *program*$_4$, vt: 'To enumerate a series of steps a device must carry out to accomplish a specific task'.

Furthermore, if a query is formulated using a wildcard (*program**) other related terms can be retrieved as shown in Table 3.3.

Table 3.3 Concordances of *program**

components available. Writing a	program	that works perfectly with every combina
cated. It is not uncommon for a	programmable	keyboard, in which some keys have spec
indows. Windows is not a single	programme	As such, but more a shell. You may have
gramming language and it can be	programmed	to do things like modify files and send
and show you how to become a C	programmer	starting at the beginning. You will be
utomated manner (at least for a	programming	environment) as experience is gained.
IOS Data Area from base memory.	Programming	the memory hole or any kind of implement

These new linguistic items are related formally as well as semantically to one the meanings of *program* explained above.

> *program*$_1$, *n., programme; program*$_3$, *vt; programming; programmer*
> *program*$_2$, *n., program*$_4$, *vt; programmable*

Corpus sentences contain linguistic units that share semantic relations with *program* which differ from those mentioned earlier: *memory* (*to program a memory*); *environment* (*a programming environment*); *write* (*write a program*); *keyboard* (*programmable keyboard*). The corpus should implicitly or explicitly reveal that *program* shares many kinds of relations with other terms as shown in Table 3.4.

Table 3.4 Relations shared by *program* with other terms in the field of computing

program$_1$, n.	A set of instructions written in a specific language …	*computer ~, code, analyzer, script, routine, macro, virus, worm, write a ~, develop a ~, to program, programming, compile a ~, debug a ~, the ~ runs, the ~ terminates*, etc.
program$_2$, n.	A series of programs designed to carry out specific tasks defined …	*word processor, software ~, application, spreadsheet, antivirus, spyware, software, browser, driver, firewall, freeware, download a ~, install a ~, uninstall a ~, load a ~, run a ~, quit a ~, the ~ crashes*, etc.
program$_3$, *vt*	To write a series of instructions using a computer language in order to have a computer carry out a task or a series of tasks automatically.	*code, develop, write, programming, programmer*, etc.
program$_4$, *vt*	To enumerate a series of steps a device must carry out to accomplish a specific task.	*configure, programming, programmable, reprogram*, etc.

Figure 3.2 depicts this second type of analysis graphically. Terms are identified, they are then observed in linguistic environments to understand their meaning. Various kinds of relations can be found between these terms and others. We can compare Figure 3.2 to Figure 3.1 that illustrates a knowledge-driven analysis.

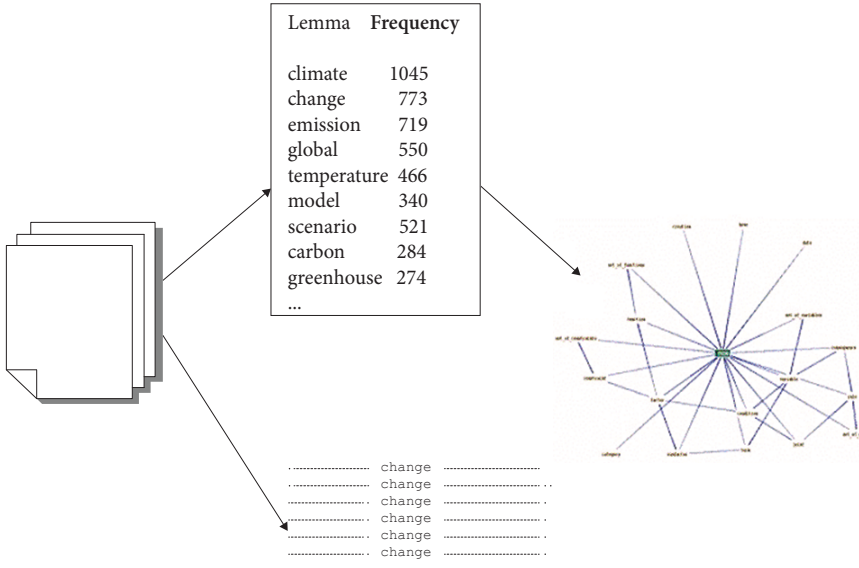


Figure 3.2 A schematic representation of an approach to terms driven by the lexicon used in specialized texts

The next sections further explore other issues that this basic comparison involves.

3.1.1 Meaning versus concept

Any terminological analysis (regardless of the approach) must inevitably deal with ***linguistic content*** that is called ***sense***, ***meaning*** or ***concept***. There are different ways to approach linguistic content (Kleiber 1999; Geeraerts 2010; Cruse 2011) and our short comparison highlighted some of them.

The General Theory of Terminology (GTT) and many other approaches concerned with knowledge representation have their own perspective on linguistic content. Traditionally, terminology uses the term *concept* rather than *meaning* when referring to the items of knowledge in specialized domains. The '**concept**' in this context represents objects of the real world: it is assumed to be an abstract generalization of the characteristics that these objects share. In other words, a

concept corresponds to a ***category***.[9] A linguistic expression is linked to a concept and labels it. As seen in Section 2.1, concepts come before linguistic labels in terminological analysis (according to an ***onomasiological approach***). This leads terminologists to consider that different kinds of linguistic expressions (lexical units, multiword expressions, phrases) and non-linguistic expressions (symbols) can be relevant terms. Hence, the focal point consists in strictly delimiting concepts; linguistic labels come afterwards and can vary quite drastically in form.

In terminology, ***concepts*** correspond to items of knowledge in specialized fields and are defined according to criteria on which experts agree (as in the animal species example given in Section 2.1). A list of ***necessary and sufficient conditions*** determine if an object can belong to a category. In other words, objects belong to the category if they meet all the conditions and meet those conditions only; conversely, objects that lack one condition or that have additional features cannot belong to the category.[10] Concepts can also be redefined when new knowledge is available. In some cases, the definitions of concepts are not yet consensual and experts are still in the process of discussing them.

This first ***approach*** to linguistic content or meaning is ***knowledge-driven***. Figure 3.3 illustrates the different parameters of a knowledge-driven approach when considering the relationship between the concept and linguistic expressions used to label it.

9. This view is also held to a certain extent in linguistics where the meaning of linguistic units is often explained in terms of concepts: "concepts are the main constituents of sense, and [that] sense (and hence concepts) constrains (even if it does not completely determine) reference." (Cruse 2011: 46) Concepts are associated with representations stored in humans' minds. However, many assume that this mental representation corresponds to the perception speakers have of objects in the world (or different worlds, since some concepts are imaginary) and this differs from the view held in traditional terminology. Human beings are exposed to objects of the world and to social conventions and they construe concepts accordingly. There are several reasons to believe that this mental representation is shared to a large extent by different human beings. One of them is that human beings can communicate.

10. During centuries, concepts were explained against a checklist of necessary and sufficient conditions. Work in cognitive science has challenged this assumption and explains the 'concept' in terms of a prototype (Rosch 1978). We will come back to this in Chapter 5.

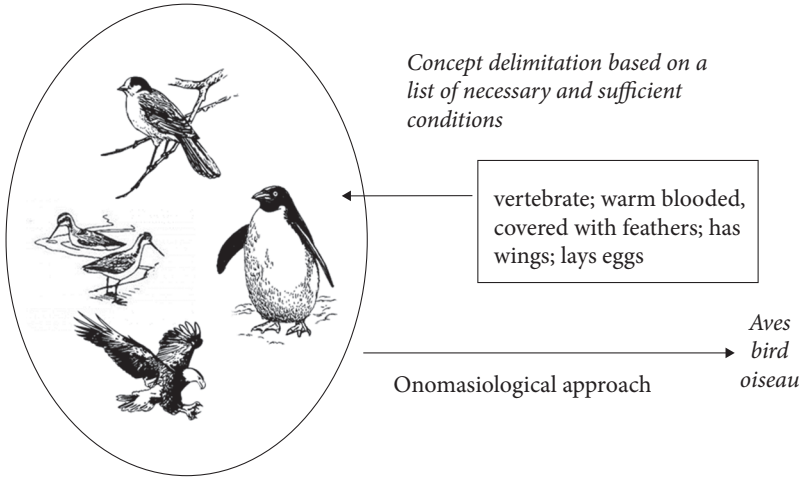


Figure 3.3 Concepts and linguistic expressions in a knowledge-driven approach

This standard view defines 'concepts' differently than other perspectives. For instance, cognitive science seeks to understand how concepts are stored and organized in the human mind. Concepts defined by experts may or may not coincide with the way meaning is represented in non-experts' minds. For instance, in zoology, a bird is defined as an animal that is warm-blooded, has feathers, and lays eggs. It differs from a reptile according to its morphological properties and from mammals according to the way it reproduces. We can probably assume that in speakers' minds, the ability to fly is an extremely important feature associated with 'bird'. In zoology, this property is irrelevant for defining the category 'bird', and this allows zoologists to classify ostriches as birds.

There are a number of problems associated with the knowledge-driven approach when considering data in running text. The most important one is that not all concepts in all special subject fields are defined by experts according to a set list of characteristics. In fact, there are still many cases where concepts are not delineated as clearly as species in zoology. Even in zoology, some species still raise classifications issues. The second problem is that the knowledge-driven approach is first interested in concepts; linguistic labels for these concepts are only considered afterwards (after the concept has been properly delineated and/or defined).

A quite different approach to linguistic content defines the ***meaning*** of a linguistic expression based on its interactions with other linguistic units. This approach assumes that ***language*** is a ***system***, i.e. a structured whole (de Saussure 1916). According to this paradigm called ***structuralism***, a linguistic unit is not

defined on the basis of categories that are established in the real world.[11] Rather, it is considered according to the place it has in the language system. Furthermore, meaning is strictly linked to a linguistic expression and neither can be analyzed independently.

The second approach no longer depends on a prior delimitation of concepts. Rather, it starts with linguistic expressions and then identifies the meanings that are associated with them (see Figure 3.4). In other words, the ***approach is semasiological***. Since terminology is concerned with linguistic units that convey specialized meaning, a relevant starting point is the lexical unit (Section 4.3). We will thus refer to this approach as the ***lexicon-driven approach***.

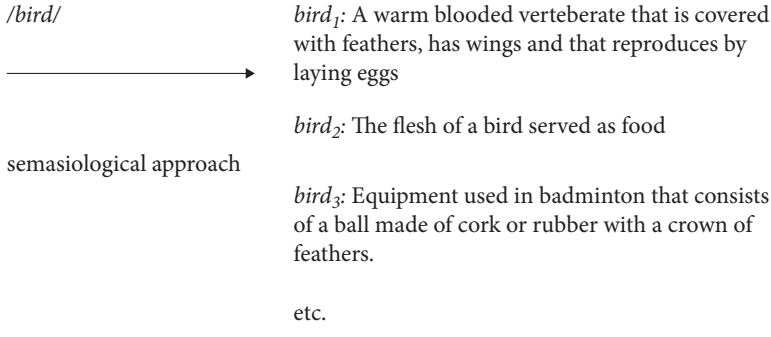


Figure 3.4 Lexicon-driven approach to the meaning of linguistic expressions

An important strand of this approach that will be called for the time being ***relational***[12] (Cruse 1986; Lyons 1977; Meľčuk et al. 1995) consists in delimiting meanings of lexical units based on the relations they share with other units. Several relations can be observed in language: two lexical units can convey the same meaning (*start* ↔ *begin*); the meaning of a lexical unit can include that of another (*flower* ↔ *tulip*); two lexical units can carry opposite meanings (*sustainable* ↔ *unsustainable*), and so on. These relations can be used to support meaning distinctions. For instance, consider the adjective *clean* in the following sentences:

11. By stating this, we are not saying that there is no connection whatsoever between language and reality. Rather, we stress the fact that this is not how the meaning of linguistic expressions is considered in the discussion that follows. This being said, the possible connections between language and the semantic and pragmatic phenomena it expresses (including our construal of reality) is a topic debated in different circles, especially in cognitive linguistics (Croft & Cruse 2004).

12. This terminology is based on Geeraerts (2010).

However, it is not a CLEAN *fuel, releasing more sulphur dioxide than gas.*
**CLEAN *hands protect against infection. Protect yourself.*

In the first sentence, the meaning of *clean* can be connected to that of other words, such as *green*, and is opposed to *polluting* or *fossil*, whereas, in the second sentence, *clean* can be linked to *immaculate* and opposed to *dirty*. The second occurrence can also be related to a verb, i.e. *to clean*, while this is not possible for the first occurrence. This evidence allows us to claim that *clean* has at least two different meanings.

The merit of the relational approach lies in the fact that it provides useful criteria when considering meanings attached to linguistic units, and more specifically, lexical units. As we shall see further on, the approach is of great assistance in terminological analysis. Furthermore, it is compatible with the idea that ***structure*** is a fundamental notion in terminology. However, when looking at structure from a lexicon-driven perspective, the focus is placed on the meaning of terms and not on concepts as abstract generalizations of items of knowledge (we will see why in Chapters 7 and 8). Or course, this approach is not devoid of shortcomings. One of them is that, when applied unilaterally, it does not inform us of possible connections between meaning and the real world. Another shortcoming is that it does not directly inform us on how meaning is encoded in the human mind.

In this book, it is assumed that ***terminological analysis*** must incorporate aspects of both ***knowledge-driven*** and ***lexicon-driven approaches*** to linguistic content. First, terminologists must take into account the definitions formulated by experts in the fields to which terms belong (that is, of course, when these definitions exist, which is not always the case). This perspective is taken in order to identify relevant terms in a given domain. To do this, a link between knowledge and units that express it must be established since these units should be the ones terminologists collect, analyze and describe. Additionally, the meaning of terms is considered from the point of view of a single subject field. This allows us to consider that *warming* in the field of climate change involves the atmosphere, the Earth, the Earth's surface, and so on, whereas, generally speaking, *warming* can apply to many more situations. Hence, meaning in terminology is necessarily studied from the perspective of a specific subject field and should be considered within the boundaries of that field.

Secondly, terminological analysis needs criteria and a methodological apparatus to answer many other questions that a ***lexicon-driven approach*** can provide. The lexicon-driven approach presented in this book is based on principles of the relational perspective on meaning described earlier. It allows us to handle units that are seldom taken into account by traditional terminology (i.e. verbs,

adjectives; predicative units) with adequate criteria. It can also handle various relations between terms that are often neglected in knowledge-driven perspectives. Other aspects of terms that the lexicon-driven approach considers are listed further in this section.

3.1.2 Dealing with lexical units that belong to different parts of speech

Section 3.1 showed that terms can belong to different ***parts of speech***. An analysis carried out within a knowledge-driven approach inevitably favors nouns. Consider the 'mouse' example again (Figure 2.3). The concepts identified in the small conceptual structure are all designated by nouns or noun phrases (*mouse*, *wireless device*, *peripheral device*, etc.). A typical activity associated with the concept of 'mouse' is 'to click' (*someone uses a mouse to* CLICK *on a graphical object on screen*). Is 'click' a relevant concept in the field of computing? If it is, how should it be represented in the conceptual structure? We would need to define new conceptual categories for activities and list the typical activities for which peripheral devices are designed. Additionally, a new conceptual relation should be defined to link these concepts to entities. In this case, <has-function> would be an option. Figure 3.5 shows how activity concepts could be added to a conceptual structure. In this example, verb labels were kept, but the use of nouns instead of verbs has no effect on the structure itself.

We can examine this question from a different perspective. If we asked experts and laypeople the following question: do you think *click* is a term in the field of computing in the following sentence "CLICK on the icon to open the file," the answer would probably be a unanimous *yes*. Other units mentioned ealier, *to program, programmable*, would probably elicit the same kind of answer.

However, the meaning of ***verbs*** and ***adjectives*** (and also adverbs, although adverbs rarely qualify as terminological units) cannot be fully captured in a conceptual structure. In fact, their meaning cannot be described the same way as nouns, specifically nouns that denote entities. Verbs, adjectives and adverbs (and in fact many nouns) are ***predicative units*** and require ***arguments***. *Program*$_3$, for instance, has three arguments (**X** *programs* **Y** *in* **Z**); *click* also requires three arguments (**X** *clicks on* **Y** *with* **Z**)).[13] If these units are to be considered in terminological analysis, lexicon-driven approaches provide useful methods to describe them.

13. Chapter 6 presents the notions of 'predicative unit' and 'argument'.

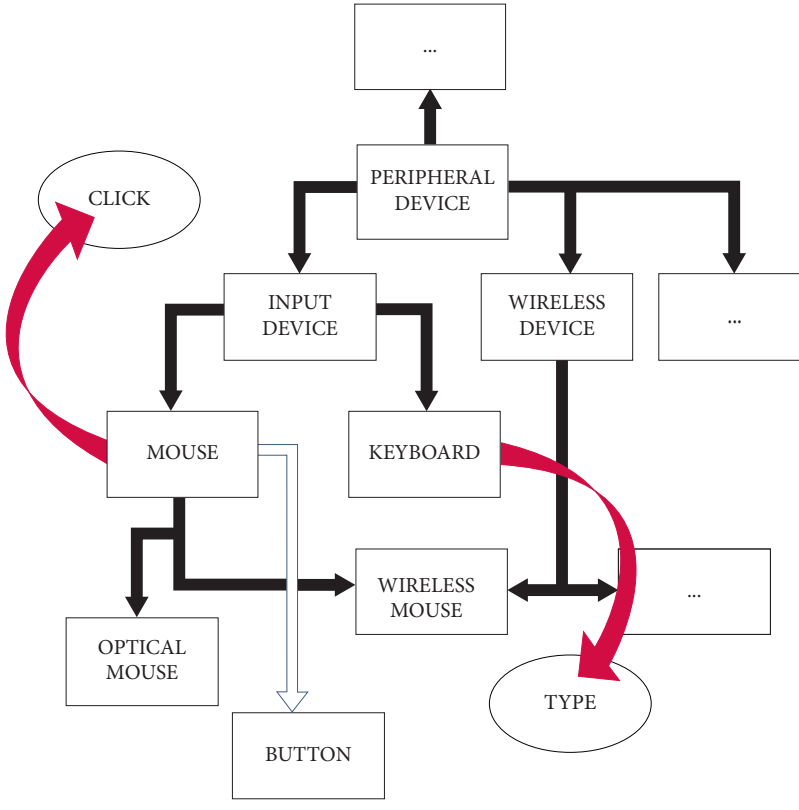


Figure 3.5 A conceptual structure with activities

3.1.3 Making (fine-grained) semantic distinctions

Lexical items can convey multiple meanings, some of which are relevant in the same domain. Examples were given in the field of computing. As a noun, *address* can denote: 1. 'the unique identifier of a user', or 2. 'an identifier of a location on a storage medium'. *Program* as a noun can be defined as: 1. 'A set of instructions written in a computer language …,' or 2. 'A series of programs designed to carry out specific tasks …' As a verb, *program* designates: 3. 'To write a series of instructions using a specific language in order to have a computer carry out a task or a series of tasks automatically', or 4. 'To enumerate a series of steps a device must carry out to accomplish a specific task'.

Many more examples can be found in other fields of knowledge. In an environment corpus, *emission* can designate 'the release of a substance into the environment' (*scientists have demonstrated that the sustained* EMISSION *of greenhouse gases will inevitably lead to global warming*). It can also refer to the substance that is released: 'a discharge of pollutants, such as greenhouse gases, in the environment' (*reduce* CO_2 EMISSIONS *by approximately 750 tonnes*). Similarly, *environmental*

can mean: 1. 'that concerns the environment' (as in *environmental impact*) or 2. 'that is designed to have minimum impact on the environment' (as in *environmental policy*).

Consider the following example. The French item *terre* has up to four meanings, all of which are relevant for the environment:

> *Terre* 1, n. f. (En the Earth). 'a planet of the solar system inhabited by living organisms'. (*la Terre est le cadre de variations climatiques*)
> *terre* 2, n. f. (En land). 'a surface of the Earth occupied by continents or islands and not covered by water'. (*Expression des forces naturelles qui agitent les éléments – la terre, la mer, l'atmosphère*)
> *terre* 3, n. f. (En land). 'an area of ground used for specific purposes, e.g. farming purposes'. (*conversion de terres forestières en terres agricoles*)
> *terre* 4, n. f. (En earth, soil). 'the substance in which plants grow'. (*couche de terre végétale*)

Terminologists will need to support these ***semantic distinctions*** with criteria. A lexicon-driven approach such as the one briefly described in Section 3.1.1 can examine the relations of each meaning conveyed by polysemous items using the meanings of other lexical units. Table 3.5 shows how the meanings of the items in this section are linked to different lexical units.[14]

To support the same distinctions with a knowledge-driven approach, we would need to associate each meaning with a specific knowledge item in a conceptual structure. However, this appears to be a rather circular venture. Since knowledge structures are defined according to criteria established by experts (or suggested by terminologists when these structures are lacking), they necessarily reflect a given perspective on knowledge and indirectly on linguistic expressions chosen to express this knowledge.

Table 3.5 Polysemous lexical items and related lexical units

Polysemous item	Definition or short explanation	Related lexical units
address$_{N\,1}$	'The unique identifier of a user'	*email address, login, user name, password, at sign*
address$_{N\,2}$	'An identifier of a location on a storage medium'	*addressing, to address, memory address*
program$_{N\,1}$	'A set of instructions written in a specific language …'	*computer program, to program, programming, loop, routine, language*

14. More will be said on how meaning distinctions can be made in Chapter 5.

Table 3.5 (*Continued*)

Polysemous item	Definition or short explanation	Related lexical units
program$_{N\,2}$	'A series of programs designed to carry out specific tasks defined …'	*programmable, reprogrammable*
program$_{V\,3}$	'To write a series of instructions using a computer language in order to have a computer carry out a task or a series of tasks automatically'	*write, programming, language, develop, debug, compile*
program$_{V\,4}$	'To enumerate a series of steps a device must carry out to accomplish a specific task'	*reprogram, programmable, reprogrammable*
emission$_1$	'The release of a substance into the environment'	*to release, to emit, to absorb, absorption*
emission$_2$	'A discharge of pollutants, such as greenhouse gases, in the environment'	*pollutant, greenhouse gas*
environmental$_1$	'That concerns the environment'	*ecological, environment*
environmental$_2$	'That is designed to have minimum impact on the environment'	*green, clean, polluting*
Terre$_1$	'One of the planets of the solar system inhabited by living organisms'	*Mars, Vénus, système solaire* (En *solar system*)
terre$_2$	'A surface of the Earth occupied by continents or islands and not covered with water'	*océan* (En *ocean*), *mer* (En *sea*)
terre$_3$	'An area of ground used by humans for specific purposes, e.g. farming purposes'	*terrain* (En *land*), cultiver une ~ (En *cultivate a ~*)
terre$_4$	'The substance in which plants grow'	*eau* (En water), *roche* (En rock)

3.1.4 Taking into consideration relations between terms

We saw above that *program* is semantically related to other linguistic units. The ***relations*** held by *program* with other terms can be explained as follows:

A PROGRAM$_1$ *is a kind of software.*
A bridge program (a client program | a backup program) is a kind of PROGRAM$_1$.
A routine (a loop | an instruction) is a part of a PROGRAM$_1$.
PROGRAM$_3$ means create a PROGRAM$_1$ or a PROGRAM$_2$.
PROGRAM$_1$ is a part of PROGRAM$_2$.
Programmable means that something can be PROGRAMMED$_3$.
A *programmer* is someone whose main activity is to PROGRAM$_3$.
Programming is the activity carried out by someone who PROGRAMS$_3$.

> To *write* a PROGRAM$_1$ means to create it using a specific language.
> An *environment* is a set of tools that can be used by someone whose task is PROGRAMING.
> To PROGRAM$_4$ a *memory* consists in specifying specific instructions to state the steps that must be followed to carry out a task.
> And so on.

Traditionally, terminology has been interested in the relations paraphrased in the first three statements (taxonomic and partitive). However, *program* shares many other relations with other units (typical activity, creation, typical agent, etc.). This observation applies to many other terms, of course. The examples above also show that ***relations*** between terms can be ***paradigmatic*** (in the lexicon or at a specific point in a sentence), or ***syntagmatic*** (between words that co-occur in the same sentence). Knowledge-driven approaches are usually interested in a set of relations that belong to the first category if those labels were used. If we want to take all relations into account, a more flexible system must be sought.

3.1.5 Considering the combinatorics of terms

Some of the ***relations*** identified above concern combinations of *program* with other units in sentences, i.e. they are ***syntagmatic***.

> To *write* a PROGRAM$_1$ means to create it using a specific language.
> To PROGRAM$_4$ a *memory* consists in specifying instructions to state the steps that must be followed to carry out a task.

A specific set of ***combinations*** in which terms can be found, i.e. ***collocations***, are particularly interesting in terminology and resources have started to record them (Cohen 1986; Binon et al. 2000, to name a few). Non-experts might experience problems for producing proper collocates in specialized texts. For instance, what is the typical use associated with the term *Internet* (*to browse*, *to search*)? How are the typical activities carried out by a programmer expressed (*to program, to develop, to debug, to write*)? Which terms appear in the vicinity of *degrade* in the field of the environment (*ecosystem, environment, habitat, land, region*, etc.)?

Again, the answers to these questions are more naturally supplied by lexicon-driven approaches rather than by knowledge-driven ones.

3.1.6 Considering the syntactic behavior of terms

Some terms (predicative terms) have arguments. If we consider this to be an important property for certain terms, this will trigger a need to account for another linguistic feature that has been overlooked in terminology by nearly all approaches, including recent ones, i.e. the behavior of terms in sentences. For instance, if we

are interested in the two different meanings of the verb *write* in the field of computing, we might want to know how to use them in well-formed sentences.

> write$_1$, vt: **X** writes **Y** to **Z** (as in *You can* WRITE *and rewrite to a CD-RW disc over 1,000 times*)
> write$_2$, vt: **X** writes **Y** in **Z** (as in *Java can be used to* WRITE *applications*)

Write$_1$ can be combined with its third argument through two different prepositions (i.e. *to a disk, on a disk*). On the other hand, the third argument of *write*$_2$ is linked to the verb with the preposition *in* (*in C++*). It can also be used in other syntactic positions (e.g. *Java can be used to* WRITE *programs*). This is only a small sample of the questions that one might ask when dealing with predicative units. For instance, *compatible* can be used in three typical syntactic structures (***X** is compatible with **Y**; **Y**-compatible **X**, a compatible **X***). This kind of information, however, is almost never recorded in terminological resources and is obviously not the focus of knowledge-driven approaches.

Section 3.1 presented some general differences between the knowledge-driven approach taken in traditional terminology and the lexicon-driven approach. Other differences will be highlighted and discussed further in this book. We will look at some useful frameworks (Section 3.3) that can give us a better understanding of the functioning of terms. First, Section 3.2 examines various matters regarding the basic material used in nearly all terminology projects, i.e. the corpus.

3.2 Corpus and terminology

In previous sections, many examples taken from ***specialized corpora*** were presented as evidence to support claims about terms. Terminologists and many other experts involved in terminology work almost always rely on corpora since they need to acquire knowledge in the field in which the terms or concepts they must account for are used. In fact, corpora are an integral part of most linguistic analyses as well and corpus exploration has become a standard method in lexicography. However, terminologists have relied more heavily on corpora than on other methods used by linguists, such as intuition.

In terminology, the ***corpus*** is often referred to for the following tasks:

- Acquiring knowledge in a new field of knowledge (for instance, electric transportation).
- Finding terms and attestations of terms (e.g. several occurrences of *e-bike* in different sentences and in different texts).
- Identifying different ways to express the same concept or meaning (e.g. *electric bicycle; e-bike; electric bike*).

- Spotting clues about the meaning of terms (e.g. *an electric car is a vehicle powered by electricity*).
- Finding clues about relations between terms (e.g. *an electric car is a car* expresses a taxonomic relation).
- Finding collocates of a term (e.g. *drive an electric vehicle; ride an e-bike*).

Many authors have defined criteria to compile specialized corpora for terminological purposes and devised methods to use them (Meyer & Mackintosh 1996; Ahmad & Rogers 2001; Bowker & Pearson 2002; Barrière & Agbago 2006). In addition, over the years, various tools were developed to assist terminologists with the processing of corpora: term extractors and concordancers.[15] These tools provide terminologists with a certain perspective on the data contained in corpora and indirectly on the knowledge conveyed by this data. They also draw their attention to phenomena that could otherwise be overlooked.

In this book, the corpus is also considered to be an essential part of terminological analysis since it provides the basic material to support the understanding of terms. However, we will take things a little further and resort to methods in which contexts serve to validate or invalidate possible intuitions we might have about the meaning of lexical items. This method was already used in previous pages. For instance, when looking at the adjective *clean* (Section 3.1.1), the two following sentences were given:

> *However, it is not a* CLEAN *fuel, releasing more sulphur dioxide than gas.*
> *Clean hands* protect against infection. Protect yourself.

It was mentioned that *clean* in the first sentence can be connected to one meaning of *green*, while *clean*, in the second sentence, can be linked to *immaculate*. We can now verify this assertion by trying to replace the occurrences of *clean* with the other adjectives in the same sentences. In some cases, the replacement produces sentences that are acceptable; in others, the replacement generates odd results.

> *However, it is not a* GREEN *fuel, releasing more sulphur dioxide than gas.*
> ?*However, it is not an* IMMACULATE *fuel, releasing more sulphur dioxide than gas.*
> IMMACULATE *hands protect against infection. Protect yourself.*
> ?*Green hands protect against infection. Protect yourself.*

15. More sophisticated methods and tools can also be used to extract information from corpora and structure extracted units in different ways: e.g. automated extraction of co-occurrents; distributional analysis.

Of course, the data that terminologists must handle is much more complex than these conveniently selected examples. Furthermore, it can differ quite drastically from the data lexical semanticists study since terminologists face different kinds of challenges. As was mentioned above, terminologists are seldom knowledgeable in the domain they are asked to describe. In fact, unlike lexical semanticists who can resort to their intuition and prior knowledge when analyzing most lexical items, terminologists must acquire this knowledge. They are likely to run into lexical items they never saw before or meanings with which they are unfamiliar.

This being said, it can be assumed that once part of this knowledge is acquired, the observation of terms in sentences extracted from specialized corpora allows terminologists to identify part of their ***semantic components***. In addition, differentiating terms in sentences allows them to identify relations between these terms.

Let us examine an example where contexts help confirm intuitions about a potential meaning distinction in a specialized domain. The following example is a much more realistic illustration of the sort of distinction terminologists will need to make when compared to the *clean* example. Here, it is used to determine whether *environment* has one or more than one meaning by using sentences extracted from a corpus of environment-related texts.

Wildlife diversity and abundance of all living things are determined by interactions among and between organisms and their physical ENVIRONMENTS.
(Utah State University Extension Service Department of Fisheries and Wildlife 1998)
Aplomados in coastal and tropical ENVIRONMENTS *are highly insectivorous ….*
(Campbell 1995)
The Barton Springs Salamander is clearly capable of living underground, but also inhabits surface ENVIRONMENTS. *(Campbell 1995)*

Extinctions may come about naturally through changes such as glaciation, climatic changes, lack of adaptation of species to changing ENVIRONMENTS.
(Yukon Department of Environment 2019)

Sustainable development thus involves protecting and improving the quality of the ENVIRONMENT. *(EUROPA 2007)*
Changes in climate will interact with these underlying changes in the ENVIRONMENT, *adding further stresses to a deteriorating situation.*
(Intergovernmental Panel on Climate Change (IPCC) 2006a)
By protecting the ENVIRONMENT, *the EU is not only tackling pressing problems, but also reflecting the wishes of its citizens. (EUROPA 2007)*
The large growth in emissions of greenhouse gases and other pollutants as projected in some of the six illustrative SRES scenarios for the 21st century will degrade the global ENVIRONMENT *in ways beyond climate change.*
(Intergovernmental Panel on Climate Change (IPCC) 2006)

The first four contexts show that *environment* denotes some sort of location. This is evidenced by the use of *environment* with the preposition *in* and with the verb *inhabit*. The contexts also show that *environment* refers to locations that vary and can be found in different areas (*coastal*, *tropical*, *surface*, *changing* and use of *environment* in the plural form). Finally, they reveal that the location can be used by different kinds of species (*Aplomadoes*, *organisms and their* ENVIRONMENTS, *Barton Springs Salamander inhabits* ENVIRONMENTS, *adaptation of species to surface* ENVIRONMENTS).

The following four contexts show occurrences of *environment* that have a broader and more abstract denotation. First, *environment* is used in the singular form and with the definite article. *Environment* applies to a more general reality (this is emphasized by the use of *global* in the last context). Then, it is modified by linguistic units that convey the idea of change and potentially bad change (*changes in the* ENVIRONMENT, *degrade the global* ENVIRONMENT) and others that convey the idea of a requirement for protection (*protecting and improving the quality of the* ENVIRONMENT, *protecting the* ENVIRONMENT).

This short discussion should convince most readers that we are dealing with two different meanings for *environment: environment*$_1$ and *environment*$_2$. However, we can take this a step forward and provide more corpus data to support the distinction.

Other terms are linked to *environment: habitat*, *site*, *territory, biome* and *ecosystem*. The following sentences contain attestations of the first three terms.

> *… the* HABITATS *of many species will move poleward or upward from their current locations. (Intergovernmental Panel on Climate Change (IPCC) 2006c)*
> *protecting plants and animals, and their* HABITATS. *(EUROPA 2007)*
> *Deterioration of marine* HABITATS *on the other side of the world affects our food supplies. (EUROPA 2007)*

> *The Endangered Species Act of 1973 protects land that satisfies spatial and nutritional needs and includes* SITES *for breeding, reproduction, and shelter. (Utah State University Extension Service Department of Fisheries and Wildlife 1998)*
> *For example, pregnant females seek basking* SITES *protected by thick vegetation. (New Mexico Department of Game and Fish 2017)*
> *… areas known to be used as* SITES *for relocated tortoises. (Mississippi Museum of Natural Science 2014)*

> *Eagles often have one or more alternative nests within their* TERRITORIES. *(Campbell 1995)*
> *…* TERRITORIES *of wolves in other regions of North America range from 25 to over 5,000 square miles. (New Mexico Department of Game and Fish 2017)*
> *Broods generally remain on the nesting* TERRITORY, *expanding their movements as they mature or are disturbed. (Campbell 1995)*

All terms share the following with *environment*:

- They all designate specific kinds of locations.
- These locations are inhabited by species: this is shown in phrases such as *HABITAT of many species, SITES for relocated tortoise; their TERRITORIES*.
- These locations are themselves situated in larger areas as highlighted by some modifiers of the terms: *marine HABITAT, TERRITORIES in regions*. No evidence is provided in the examples for *site*.

However, the contexts also show that:

- SITE and TERRITORY differ slightly from *habitat* and *environment* since they are combined with linguistic expressions that denote activities (*nesting TERRITORY, basking SITES, SITES for breeding, reproduction, and shelter*). This indicates that the activity for which these locations are used is an important semantic component for *site* and *territory* while it is not for *environment* and *habitat*.

The above discussion only makes sense if we consider some of the occurrences of *environment* (those that were associated with *environment*$_1$). *Environment*$_1$ is closely related to *habitat* and still closely, but a little less, to *site* and *territory*.

Looking at contexts in which *ecosystem* and *biome* appear reveals that they share a relation with *environment*$_2$: i.e. they are meronyms of *environment*$_2$. It would be odd to consider them from the point of view of *environment*$_1$.

> *Many point to technological innovation and adaptive behavior as a means for managing the global ECOSYSTEM.* *(Schwartz & Randall 2003)*
> *These models usually deal with ECOSYSTEMS or BIOMES [i.e., the collection of ECOSYSTEMS within a particular climatic zone with similar structure but differing species (e.g. the "temperate forest biome")].*
> *(Intergovernmental Panel on Climate Change (IPCC) 2006c)*

We could also look at another term, the adjective *environmental* that is connected to *environment*$_2$ but not to *environment*$_1$.

> *The EU's objective is to reduce the ENVIRONMENTAL impact of resource use.* *(EUROPA 2007)*
>
> *Impact on the environment*$_2$
>
> *The Government of Canada maintains that climate change is one of the most significant ENVIRONMENTAL and sustainable development challenges facing the globe.* *(Environment Canada 2007)*
>
> *Challenges to the environment*$_2$

It should be clear by now that contexts are extremely useful to support terminological analysis and they will consistently be used in the following pages. Of course, when describing terms in specialized resources, a more thorough analysis is required. For instance, although there are several clues that back up two distinct meanings for the noun *environment*, nothing in the contexts clearly states what *environment*$_2$ denotes. Similarly, although we can see that *environment*$_1$ shares semantic components with *habitat*, *site* and *territory*, the precise relations it holds with each of these terms are not labeled precisely.

It must also be kept in mind that some answers to questions raised by terminological analysis will not necessarily be found in contexts extracted from a corpus, even a carefully crafted one. There are a number of shortcomings that will require the judgment of terminologists and experts. Firstly, corpora do not contain all the information we need to know about a domain. Even if specialized corpora provide valuable clues about the meaning of terms, some important details will always be missing. Secondly, corpora might contain contradictions as experts do not always agree on everything. Some might even contain errors. Additionally, statements that appear in corpora can be irrelevant to terminological analysis. For instance, simple attestations of terms can be uninformative. Finally, even if some relations can be formally stated (e.g. *an electric car is a car*), it might not be the case for others (many other kinds of cars can be mentioned in a corpus, but it is unlikely that they will always be explicitly linked to *car* the way *electric car* is).

Hence, even if the corpus is a prerequisite to terminological analysis, the information it provides is seldom enough. Many questions to which there is no concrete answer can be solved by the terminologists' knowledge of specific phenomena in their language and with the help of lexical semantics methods. Additionally, even if terminologists can go a long way with the information a corpus can provide, there will always be a point where ***experts*** in the field need to be consulted. The challenge for terminologists is to find the right balance between the information supplied by corpora, the knowledge they acquire and other sources of information. They cannot rely exclusively on corpora to make decisions about terms, nor can they hide behind corpus evidence when other sources of information are lacking.

One last thing worth mentioning in relation to the use of corpora in terminology is that, unlike lexicographers and linguists who run queries on very large corpora that should ideally contain texts of different natures, terminologists use topic-specific corpora. While this is essential in order to analyze terms in a specialized linguistic environment and see how experts use them, it can also get terminologists to consider lexical items from a narrower perspective. This inevitably has an impact on the way meanings are described in terminological resources and lead to phenomena called *multidimensionality* and *meaning modulations* that are presented in Chapter 5.

3.3 Some relevant frameworks for terminology

This book argues that in order to make sense of the data supplied by specialized corpora terminologists will turn to lexical semantics frameworks. Different frameworks have been applied to the analysis of specialized lexica: ***Classes d'objets*** (lit. object classes) (Gross 1994), ***Corpus Pattern Analysis*** (Hanks and Pustejovsky 2005), ***Distributional Semantics*** (Harris 1968),[16] ***Explanatory Combinatorial Lexicology*** (Mel'čuk et al. 1995, 1994–1999), ***Frame Semantics*** (Fillmore 1976, 1982; Fillmore & Baker 2010), the ***Generative Lexicon*** (Pustejovsky 1995), among others.

In addition to these models, recent terminology work has been strongly influenced theoretically and methodologically by other strands that study the meaning of lexical units, i.e. lexicography, corpus linguistics, and cognitive linguistics. Large lexical projects, such as WordNet (Fellbaum 1998), have also had an impact on several fields, including terminology.

This book mainly refers to two frameworks, namely Explanatory Combinatorial Lexicology (ECL) and Frame Semantics (FS). They answer questions regarding the linguistic properties of terms and help address some of the issues that were raised in previous pages. Furthermore, both frameworks have led to the development of lexical resources. This following sections provide a brief description of the two frameworks based on the literature and some entries in the lexical resources derived from them.

3.3.1 Explanatory Combinatorial Lexicology

Explanatory Combinatorial Lexicology (***ECL***) (Mel'čuk et al. 1984–1999, 1995) is the lexical component of a full-fledged linguistic model called the *Meaning-Text Theory* (MTT). The origins of the theory can be traced back to the 1960s when Mel'čuk and his collaborators were involved in a project on machine translation in the former Soviet Union. Various researchers throughout the world have been working ever since on different components of language (morphology, syntax, semantics, communicative structure, etc.) and providing a linguistic modeling of these aspects deeply rooted in a structuralist tradition (labeled by Geerearts (2010) as *neostructuralist*).

ECL aims to model the numerous and complex properties of the ***lexicon***. It provides a complete system to account for all the linguistic properties of lexical units (LUs) since it postulates that the lexicon is the core element of the semantic module of a language.

16. This approach has been used chiefly in computational terminology.

> Most current linguistic theories view a description of a language as a grammar; a lexicon is taken to be an indispensable, but somehow less interesting, appendix to this grammar, where all the idiosyncrasies and irregularities that cannot be successfully covered by the grammar are stored. In sharp contrast, MTT considers the lexicon as the central, pivotal component of a linguistic description; the grammar is no more than a set of useful generalizations over the lexicon and, thus, secondary to it. (Mel'čuk 2013: 262)

ECL considers that the outcome of lexicological description is a theoretical dictionary: this dictionary – the ***Explanatory Combinatorial Dictionary*** (***ECD***) (Mel'čuk et al. 1984–1999) – specifies the syntactic, semantic and combinatorial properties of LUs. Hence, on the one hand, the description of LUs should tell us how to use them in discourse; on the other hand, it should inform us of the relations a lexical unit holds with other LUs in a given language.

Versions of the ECD have been implemented in different languages. In French, an online version called *Dicouèbe* gives access to more than 1,000 lexical units and 25,000 lexical relations (Jousse & Polguère 2005). Mel'čuk and Polguère (2007) is a printed popularized version of the ECD. Polguère heads the development of a large resource in which ECL descriptions are implemented in the form of a large lexical network (Polguère 2014). In Spanish, an online database called *DICE* gives access to collocations and to pedagogical activities designed to learn Spanish (Alonso Ramos 2004). The conversion of an existing English-French dictionary using lexical functions (lexical functions are introduced below) was carried out by Fontenelle (1997).

Each entry in the ECD is devoted to a ***lexical unit*** (the notion of 'lexical unit' will be explained in Section 4.3). ***LUs*** can correspond to ***lexemes*** (e.g. *apple*, *build, large, weep*) or non-compositional ***phrasemes***, i.e. ***idiomatic expressions*** (e.g. *ask for the moon*, *kick the bucket, a cat may look at a king*). The ECD pays special attention to lexical relations in the lexicon (paradigmatic relations) and the way LUs combine with others. ***Collocations***, for instance, are an important component of most entries. The meaning of a lexical unit may be connected to several other meanings. Take the verb *eat*, we can easily think of verbs that convey similar meanings, such as *devour*, *snack*, *dine*. We can also think of typical events that involve eating, such as *meal*, *dinner*, *breakfast*; or people who are eating, such as *eater*, *gourmet*; a verb denoting a different perspective in which an agent causes the eating, such as *feed*; and so on. Understanding a lexical unit entails understanding the place it has in the lexicon of a language.

More specifically, in an ECD the properties of LUs are distributed into four zones:

– The semantic zone that consists of a definition and connotations of the lexical unit (LU).

- The phonological/graphemic zone that specifies the spelling and/or the pronunciation of the LU.
- The co-occurrence zone that specifies the morphological, syntactic, lexical, stylistic and pragmatic co-occurrence of an LU.
- The illustrative zone that contains linguistic illustrations of an LU.

English examples will serve to illustrate part of the contents of an entry. The focus is placed on those parts to which reference is made later in the book: definition, government pattern and lexical relations.[17]

- The ***definition*** corresponds to the semantic decomposition of the meaning of an LU. The definition of *escape* (as in *he* ESCAPED *from prison*) takes the following form:

> **X** escapes from **Y** by way of **Z** to **W**: '|[**X** being in place $\mathbf{Y^1}$ where $\mathbf{Y^2}$ is, such that (something related to) $\mathbf{Y^2}$ threatens **X** and it is possible that **X** will not be able to move away from $\mathbf{Y^1}$ before the threat by Y^2 is realized,]| **X** succeeds in intentionally moving away from $\mathbf{Y^1}$ via **Z** to place **W** which causes1 that the threat by $\mathbf{Y^2}$ is not realized'. (Mel'čuk 2013: 281)

ECD definitions have several features:[18]

- The defined LU takes the form of a propositional phrase where all ***arguments*** (called ***actants*** in ECL) are stated. Arguments are represented using variables (**X**, **Y**, **Z**, etc.). The arguments then appear in the definition itself.[19]
- A definition must disambiguate the LU completely, i.e. it must state the semantic similarities and differences between the LU and other related ones (and potential substitutes).
- The definition also makes explicit paradigmatic semantic links between the LU and other ones (e.g. synonyms, antonyms, other meanings conveyed by the LU).
- The definition must contribute to the description of the lexical co-occurrence of an LU, restricted (i.e. collocations in which it appears) as well as

17. What follows is largely based on Mel'čuk (2013); however, some of the technical encoding was simplified.

18. A selection was made in what is in the framework a very complex and well documented topic.

19. Of course, this rule applies to specific kinds of LUs, namely predicative and quasi-predicative units. These are introduced in Chapter 6.

non-restricted (free combinations). In other words, the contents of the definition should be compatible with the co-occurrence of the LU.

– The ***government pattern*** of an LU specifies the active ***valence***, i.e. the possible combinations between the LU and the realizations of its arguments.
HELP$_1$ (as in *John will HELP you clean up the house*).

X ↔ I	Y ↔ II	Z ↔ III	W ↔ IV
1. N	1. N	1. V_{INF} 2. *to* V_{INF} 3. *with* N 4. *with* V_{GER} 5. *in* V_{GER} 6. $PREP_{DIR}$ N	1. *with* N 2. *by* V_{GER}

(Meľčuk 2013: 309)

This government pattern first states the correspondence between the arguments of HELP$_1$ (**X**, **Y**, **Z**, and **W**) at the deep syntactic level (I, II, III, and IV). At the surface syntactic level, the first and second arguments can be realized in the form of nouns (N), the third argument, in the form of an infinitive verb (V_{INF}), an infinitive verb linked to *help* by means of a preposition (*to* V_{INF}), a noun linked to the verb *help* by means of a preposition (*with* N), and so on.

– The ***lexical relations*** zone lists the set of relations shared by an LU with other lexical units. Lexical relations include what is called in this framework *semantic derivation*[20] and restricted lexical co-occurrence, i.e. collocations. Lexical relations are represented with a system called ***lexical functions*** (***LFs***).[21] Some examples of simple LFs are reproduced below.

Syn(*bicycle*) = *bike* (exact synonym)
Gener(*car*) = *vehicle* (generic LU)
Anti(*love*) = *hate* (antonym)
Contr(*sea*) = *land* (contrastive)
S_1(*teach*) = *teacher* (the typical first argument, in this case the **Agent**)

20. In ECL, *semantic derivation* is preferred over *paradigmatic relation*.

21. Lexical functions will be further detailed in Chapter 8 (Section 8.2.1.1). For the time being, we can say that it is a semi-formal system that accounts for recurrent lexical relationships and general and abstract recurrent meanings.

S_2(*teach*) = *subject* (the typical second argument, in this case the **Material**)
S_3(*teach*) = *student* (the typical third argument, in this case the **Recipient**)
Sloc(*teach*) = *class* (the typical location when 'teach' takes place)
Able$_2$(*program*) = *programmable* (a second argument that can be + 'program')
Magn(*storm*) = *severe* (intensifier)
Fact$_2$(*teacher*) = *teach* (the inherent activity related to teacher when teacher is subject)
Real$_1$(*Internet*) = *browse* (the typical activity carried out by the first argument of 'Internet' when *Internet* is first complement)

In terminology, ECL is especially helpful to account for the meanings of terms, describe the argument structure of predicative terms (as well as quasi-predicative ones), make semantic distinctions, formalize definitions, and describe the syntactic behavior of terms. Above all, it provides us with the most sophisticated system to represent relations between terms (paradigmatic as well as syntagmatic).

To my knowledge, the first author who suggested applying ECL to the lexical expressions found in specialized fields is Frawley (1988). The author argued that the purpose of a specialized dictionary is to provide a precise characterization of the lexicon and that ECL could be used to support this kind of work. Other authors such as Gentilhomme (1994), Grass (1999), Binon et al. (2000), Marcel (2000), and L'Homme (2007, 2012) also refer to ECL and emphasize its capacity to handle some properties of the specialized lexicon that knowledge-driven frameworks cannot.

ECL has been applied in different fields of knowledge, such as mathematics, biology, law, environment, computing. However, to this day, there are few full-fledged applications to the specialized lexicon. Binon et al. (2000) applied it to a business dictionary. My research group used it to describe terms in two resources: a terminological database on computing and the Internet (*DiCoInfo. Dictionnaire fondamental de l'informatique* 2018) and another one on the environment (*DiCoEnviro. Dictionnaire fondamental de l'environnement* 2018). Some authors have studied the possible combination of general language resources with specialized ones that apply the principles of ECL (Ingrosso & Polguère 2015; L'Homme & Polguère 2008).

3.3.2 Frame Semantics

Frame Semantics*, *FS (Fillmore 1976, 1982, 1985; Fillmore & Baker 2010) is a framework based on cognitive assumptions and aims to provide a theoretical explanation for the relationship between language and how human beings represent situations in their minds. As Fillmore stated in his early work on Frame Semantics:

> A particularly important notion, figuring especially in recent work in linguistics, cognitive psychology and artificial intelligence, in the notion that goes by such names as "frame", "schema" and "scenario." Briefly, the idea is that people have in memory an inventory of schemata for structuring, classifying, and interpreting experiences, and that they have various ways of accessing these schemata and various procedures for performing operations on them. Some of these schemata may be psychologically built in (such as the various aspects of the body schema, the identity of the local hues in the color spectrum, and perhaps what the gestalt psychologists call "good figures" – see Rosch[22]), others may owe their existence to perceived cause-effect relationships in the world, while still others may depend for their existence on symbolization.
>
> The concept of frame does not depend on language, but as applied to language processing, the notion figures in the following way. Particular words or speech formulas, or particular grammatical choices, are associated in memory with particular frames, in such a way that exposure to a particular linguistic form in an appropriate context activates in the perceiver's mind a particular frame – activation of the frame, by turn, enhancing access to the other linguistic material that is associated with the frame. (Fillmore 1976: 25)

As stated by Fillmore the notion of 'frame' covers much more than the lexicon, but given the topic of this book, we will focus on the applications of the framework to lexical units. We will often refer to FrameNet (2018), the linguistic resource that applies the theoretical principles postulated by Frame Semantics to the lexicon.

Frame Semantics (FS) assumes that the meanings of ***lexical units*** are construed against a background of experience, beliefs or practices that are based at least partly on social and cultural institutions. Several examples have been given in the literature on FS to support the assumption to the effect that our understanding of lexical units involves a larger background, a broader situation that comprises participants and presuppositions. For instance, Fillmore argues that the difference between *ground* and *land:*

> … appears to be expressed by saying that LAND designates the dry surface of the earth as it is distinct from the SEA, whereas GROUND designates the dry surface of the earth as it is distinct from the AIR above it. The words "land" and "ground", then, differ not so much in what it is that they can be used to identify, but in how they situate that thing in a larger frame. (1982: 121)

We can further explain the notions of 'background knowledge', 'frame' and 'participant' with an everyday example. We all have ***background knowledge*** of a situation

22. This is a reference to prototype theory mentioned earlier in this chapter.

in which someone stays in a location for a period of time and where this person carries out activities associated with daily life. This situation involves two central ***participants***: someone who experiences it (an owner, a resident, a tenant), a place where people carry out activities (a city, an area, a house). The situation can also include other peripheral participants such as a specific time (in 1990, last month) or a duration (for 24 months, during 10 years). A native English speaker will probably relate LUs such as *camp, inhabit, live, occupy, reside, resident* to this situation. Native French speakers would probably think of LUs such as *occupant, habiter, occuper, occupation, résider*, *vivre*.

In Frame Semantics, this background knowledge is structured in the form of ***semantic frames***. More precisely, a frame can be defined as the schematic modeling of a prototypical ***situation*** that includes ***participants***, which constitute its ***frame elements*** (***FEs***).

Based on these principles, ***FrameNet*** (***FN***) (Fillmore & Atkins 1992; Ruppenhofer et al. 2016) is investigating the cognitive processes that underlie the construction of meaning and the connection between these processes and the linguistic behavior of lexical units. The original FrameNet focuses on English, but there is an increasing number of FrameNets developed for other languages, such as Brazilian Portuguese (Torrent & Ellworth 2013), Chinese (You & Liu 2005), French (Candito et al. 2014; Djemma 2017), Danish (Pedersen et al. 2018), German (Boas 2009), Japanese (Ohara 2009), Spanish (Subirats 2009) and Swedish (Dannélls et al. 2014). FrameNets differ from the Explanatory Combinatorial Dictionary described in the previous section in the sense that they focus on specific aspects of the linguistic behavior of lexical units and do not aim to provide an exhaustive account of their properties.

In Frame Semantics, the meanings of LUs are understood, analyzed and described according to background knowledge captured in semantic frames: LUs are said to 'evoke' a frame. In FrameNet, frames and LUs are described in separate but interrelated modules. The content of these modules is illustrated below with the 'living somewhere' situation that was informally introduced above. This situation is structured in a frame called RESIDENCE. The description provides:

- A definition of the frame along with examples of sentences containing LUs that evoke this frame.
- A list of participants, called ***frame elements FEs***.
- LUs that evoke this frame (Figure 3.6).

Frame elements (FEs) are divided into two categories: ***core FEs,*** necessary for defining the frame; and ***non-core FEs*** that are optional participants. In Frame Semantics, this distinction is crucial since core FEs are always cognitively active in a frame. In contrast, non-core FEs need to be mentioned explicitly to be activated. Both types of FEs can be further specified in terms of Semantic Types: for

instance, in the RESIDENCE frame, the Means is further specified as State-of-affairs (Figure 3.6). More examples are given to illustrate how FEs are realized in this frame.

English LUs that evoke the frame are listed along with an indication of their part of speech. The LUs that evoke the RESIDENCE frame, include verbs and nouns that denote activities or entities.

RESIDENCE

Definition: This frame has to do with people (the Residents) residing in Locations, sometimes with a Co-resident.

Peter LIVES in New York
Sue is an INHABITANT of Los Angeles.

Core FEs

Co-resident [Co-resident]	A person or group of people that the resident is staying with or among. *Boris still LIVES with his parents.*
Location [Location]	The place in which somebody resides.
Semantic Type: Location	*Sue LIVES in Berkeley.*
Resident [Res]	The individual(s) that reside at the Location. *Hannah LIVES in San Francisco.*

Non-core FEs

Circumstances [cir]	Circumstances describe the state of the world (at a particular time and place) which is specifically independent of the event itself and any of its participants.
Depictive [Dep]	The state of a Resident during their residence.
Frequency [fre]	Frequency at which event occurs.
Semantic Type: Degree	
Manner [Manr]	Manner of performing an action
Semantic Type: Manner	
Means [Means]	An act of the Resident that enables them to reside in the Location.
Semantic Type: State_of_affairs	*He DWELT in the wild for years by foraging nuts and berries and occasionally stealing.*
Time [tim]	This FE refers to the Time when the Resident resides.

Lexical units: bivouac.n, bivouac.v, camp.n, camp.v, camped.a, camper.n, dwell.v, dweller.n, inhabit.v, inhabitant.n, live.v, lodge.v, occupant.n, occupy.v, reside.v, resident.n, room.v, room-mate.n, shack up.v, squat.v, squatter.n, stay.v, tenant.n

Figure 3.6 Definition of the frame RESIDENCE and Frame Elements (FrameNet 2017)

Lexical units that evoke the RESIDENCE frame include *live*. However, *live* evokes four other frames since it is polysemous. Each meaning must be considered against different background knowledge.

> DEAD_OR_ALIVE: A Protagonist is in the dynamic, maintained state of being alive or has exited this state. This frame is also evoked by *alive.a, dead.a, dead.n, deceased.a, dirt nap.n, late.a, life.n, lifeless.a, living.a, living.n, nonliving.a, undead.a, undead.n.*
>
> THRIVING: An Entity is in a state such that it participates in a preponderance of states and events which are desirable for it. This frame is evoked by the following lexical units: *do.v, fare.v, flourish.v, languish.v, live.v, prosper.v, prosperity.n, slump.n, thrive.v.*
>
> MANNER_OF_LIFE: An Experiencer actualizes a certain pattern of behavior, a Lifestyle, which persists for a significant period of time in the Experiencer's life and is recognized as part of his character or normal routine. In addition to the verb *live*, *life*, and *lifestyle* are said to evoke this frame.
>
> LIVING_CONDITIONS: An Experiencer must cope with some Condition for a prolonged period, usually against his or her will.
>
> (FrameNet 2019)

Each LU that appears in a frame is further characterized in a separate module that accounts for the projection into syntax of the characterization given in the frame. More specifically, this module provides:

- A definition of the LU; it can be written by FrameNet linguists or retrieved from an existing dictionary (Figure 3.7).
- The ***syntactic realizations*** of the Frame Elements (Figure 3.7).
- The ***valence patterns*** of the Frame Elements (Figure 3.8).

Inhabit, v. RESIDENCE Definition: COD: live in or occupy.

Frame Elements and Their Syntactic Realizations:

Frame Element	Number Annotated	Realization(s)
Location	(26)	NP.Ext (11) NP.Obj (15)
Resident	(26)	PP[by].Dep (8) CNI.-- (3) NP.Ext (15)
Time	(1)	PP[for].Dep (1)

Figure 3.7 Syntactic realizations for the LU *inhabit* (FrameNet 2017)

Syntactic realizations list the different phrase types and syntactic functions of FEs in sentences. For instance, with the LU *inhabit*, the **Resident** can be realized as a noun phrase (NP) and as an external (Ext) (it can also be non-instantiated[23]). The ***valence patterns*** table summarizes how FEs can be combined and ordered. For instance, in some sentences containing *inhabit*, the FEs are deployed according to different kinds of patterns: **Location**, **Location** and **Resident; Location** and **Resident; Location**, **Resident** and **Resident**, and so on.

Valence Patterns:

Number Annotated	**Patterns**		
1 TOTAL	Location	Location	Resident
(1)	NP Ext	NP Ext	PP[by] Dep
20 TOTAL	Location	Resident	
(2)	NP Ext	CNI --	
(7)	NP Ext	PP[by] Dep	
(11)	NP Obj	NP Ext	
4 TOTAL	Location	Resident	Resident
(4)	NP Obj	NP Ext	NP Ext
1 TOTAL	Location	Resident	Time
(1)	NP Ext	CNI --	PP[for] Dep

Figure 3.8 Valence patterns for the LU *inhabit* (FrameNet 2017)

A third module provides ***contextual annotations*** of sentences in which LUs appear. In the English FrameNet, these sentences are extracted from the British National Corpus. Figure 3.9 shows a sample of the annotations for the verb *inhabit*.

23. In FrameNet, core frame elements that are not instantiated are annotated if they are "conceptually salient" (Ruppenhofer et al. 2016: 28). In Figure 3.7, three occurrences of the **Resident** are labeled as a constructional null instantiations (CNI). This means that they are omitted in certain constructions (a passive construction, for instance). Other null instantiations include Definite null instantiations (DNI) and Indefinite null instantiations (INI).

In 1926, the year of his birth, [2 billion people $_{Resident}$*] INHABITED [this planet* $_{Location}$*].*
[White and coloured peoples $_{Resident}$*] will INHABIT [the Earth* $_{Location}$*], as today, but with one essential difference.*
[Blacks $_{Resident}$*] (perish the thought) now INHABIT [his old house* $_{Location}$*] and pitched battles between gangs and police make the streets untenable.*
As the brothers disliked one another, [E. F. $_{Resident}$*] INHABITED [the house* $_{Location}$*] in term-time, then went abroad, leaving it to A. C. for the vacations.*
And now [they $_{Resident}$*] INHABITED [their house* $_{Location}$*].*

Figure 3.9 Annotated sentences with the LU *inhabit* (FrameNet 2017)

As can be seen from Figures 3.6 to 3.9, the connection between the conceptual level of the description (frame) and the linguistic levels (lexical entry and contextual annotations) is made explicit with a consistent labeling of lexical units evoking the frame and frame elements throughout the three modules.

Finally, frames are linked via a network of relationships defined by FrameNet lexicographers. A tool called *FrameGrapher* offers different views on relations shared by frames. Figure 3.10 shows some of the ***relations*** defined between the frame RESIDENCE and other frames. The edge between RESIDENCE and TEMPORARY_STAY (in red) represents inheritance; the link between RESIDENCE and COLONIZATION (in green) stands for 'using'. Other relationships include 'precedes' (in black) and 'perspective on' (in pink). The establishment of these relationships relies on a thorough analysis of the contents of frames.[24]

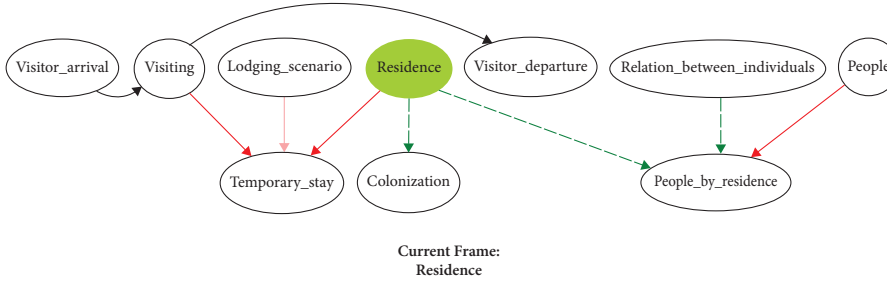


Figure 3.10 Graphical visualization of relations between the frame RESIDENCE and other ones (FrameNet 2017)

24. Relations between frames are established on the basis of frame elements within frames.

In terminology, Frame Semantics is especially helpful to account for the participants involved in the meaning of terms (especially, but not exclusively, terms denoting events and properties) and to describe the syntactic behavior of terms. Furthermore, it provides an apparatus to connect the linguistic properties of terms with abstract conceptual representations, i.e. frames. It also offers an interesting way to capture complex semantic phenomena, such as meaning modulations.

Recently, many terminologists have seen in Frame Semantics and FrameNet some potential to characterize the specialized lexicon in a way that differs from the perspectives offered by other frameworks. It even led to an approach to terminology called *Frame-based terminology* (Faber 2012, 2014). It has been applied to the fields of football (Schmidt 2009; Dicionário da Copa do Mundo 2014), Law (Pimentel 2013), medicine and biology (Dolbey et al. 2006; Wandji et al. 2013), linguistics (Malm et al. 2018), computing (Ghazzawi 2016) and the environment (Faber 2012; L'Homme 2018).

The next chapters will show how both Explanatory Combinatorial Lexicology and Frame Semantics, the principles on which they are based as well as the methods developed to compile the ECD and FrameNets can help terminologists handle terms and the different questions they raise.

Of course, these frameworks diverge in terms of assumptions and were designed to address different issues in language. One of the main differences between them stems from the fact that ECL focuses on the linguistic properties of lexical units whereas Frame Semantics aims to establish a connection between language and abstract background knowledge. Hence, FS incorporates two descriptive levels: a conceptual level that accounts for a given situation and its participants and a linguistic level that shows how this situation is instantiated in language.

However, Frame Semantics and Explanatory Combinatorial Lexicology do share some methodological principles. The 'lexical unit' is defined according to very similar criteria. Both frameworks account, albeit differently, for the interface between the semantic and the syntactic properties of LUs. The similarities between the frameworks are more easily perceptible in the lexical resources which are based on them.

It is our contention that FS and ECL complement each other. ECL allows us to provide a very detailed picture of the lexico-semantic properties of terms. FS goes slightly beyond and allows us to connect terms and their properties to broader situations. This connection informs us on how events, properties and even entities interact in a given field of knowledge.

3.4 Questions that lexical semantics cannot answer

In the previous sections, it was often emphasized that a knowledge-driven approach imposes a certain perspective on terms. This is also true for approaches

based on ***lexical semantics***. In fact, this applies to all theoretical explanations that hinge on assumptions.

Lexical semantics provides answers that cannot be supplied by a knowledge-driven approach and sheds a different light on terms, but it cannot meet the needs raised by all applications. Here are some questions for which lexical semantics (at least the models to which I refer) is not properly equipped to address. In addition, depending on the chosen lexical semantics framework, certain properties of terms are better explained than others. This section presents limitations that remain true for most frameworks.

- Since lexical semantics frameworks are descriptive rather than prescriptive, they do not supply criteria that could be useful for standardization activities. They are not designed to help select an ideal linguistic expression to label a concept in a list of competing expressions. Experts involved in these kinds of activities could take lexical criteria into account in a preliminary analysis of terms, but the final decision will have little to do with lexical semantics.
- Lexical semantics models – even the models that are interested in relations in the lexicon – are not suited to capture domain knowledge the way taxonomies, thesauri, and ontologies do. In fact, modeling knowledge requires strategies in which linguistic aspects are ignored to a large extent (see Section 2.3 where this was explained in more detail). Lexical semantics, on the other hand, focuses on the linguistic properties of lexical units. Hence, there is some degree of incompatibility between the two approaches and each is designed to capture different phenomena. For instance, explaining the 'greenhouse effect', its causes, and its consequences, is not telling us what part of speech the linguistic expression *greenhouse effect* belongs to, if it has synonyms, and with which verbs it can be combined.
- Lexical semantics can provide clues as to how humans encode meaning in their minds, but these clues are indirect. They cannot directly explain how lexical units, terms or their meanings are stored and processed by human beings. They can, however, supply hypotheses that can later be tested with experimental methods for validating or invalidating them.

Summary

Considering terms from a knowledge-driven or a lexicon-driven perspective leads to very different kinds of analyses and each raises specific questions about meaning or the types of linguistic expressions that can obtain terminological status. Although both are necessary in terminology work, lexicon-driven approaches are probably more adequate to account for the behavior of terms in running text.

In contrast, knowledge-driven approaches are suited to handle different kinds of applications in which knowledge modeling is required.

Corpora are an essential part of terminology work and are used by terminologists to acquire knowledge, familiarize themselves with terms, make subtle meaning distinctions, and establish relations between terms. However, the information supplied by corpora needs to be supported or completed by other sources of information such as experts.

Various frameworks in lexical semantics can be and have been used to analyze and describe terms. This book focuses on Explanatory Combinatorial Lexicology and Frame Semantics. However, other frameworks will be mentioned when relevant.

The next chapter examines the notion of 'term' and presents criteria to identify terms and distinguish them from other lexical units.

Further reading

The theoretical and methodological consequences of knowledge-driven versus lexicon-driven approaches to terminology are examined in L' Homme (2004a) and L'Homme and Bernier-Colborne (2012). An extremely interesting point of view on the relationship between the lexicon and knowledge representations (namely, ontologies) is presented in Hirst (2009).

A number of excellent introductions to lexical semantics can be consulted (Cruse 1986; Palmer 1976; Polguère 2016, among others). Lyons (1977) is the most prominent representative of the relational approach in English.

Corpora hold an important place in terminology work. Bowker and Pearson (2002) is a very useful textbook to learn how to manage and process specialized corpora. Corpora and their use in terminology have been investigated in various ways that are beyond the scope of this book but are nevertheless highly interesting. Here is a selection of titles: Meyer and Macintosh (1996); Ahmad and Rogers (2001); Barrière and Agbago (2006); Condamines (2005).

Meľčuk et al. (1995) is an introduction to Explanatory Combinatorial Lexicology (ECL). The four volumes of the Explanatory Combinatorial Dictionary (1984–1999) have very insightful introductions that focus on a specific component of ECL. Applications of the principles of ECL to terminology are discussed in Frawley (1988) and L'Homme (2007, 2012). Descriptive applications of ECL in specialized domains can be found in Binon et al. (2000), the DiCoEnviro (2018) and the DiCoInfo (2018).

Fontenelle (2003) edited a special issue of the *International Journal of Lexicography*, which is an excellent introduction to the principles of Frame Semantics and

to FrameNet. Fillmore (1968) was published prior to work on Frame Semantics but helps understand some of the assumptions on which it is based. The author explains his ideas with regard to Frame Semantics in other articles (Fillmore 1976, 1982, 1985). Croft and Cruse (2004) discuss Frame Semantics and situate it within the broader spectrum of cognitive linguistics.

Applications of Frame Semantics to terminology and more specifically of the methodology devised in the FrameNet project can be found in Schmidt (2009), Pimentel (2013), Ghazzawi (2016) and L'Homme (2018). Faber (2012, 2014) proposes an approach to terminology called *Frame-based Terminology* that implements principles of Frame Semantics.

Applications of other lexical semantics frameworks to terminology can be found in many publications from which a selection was made. Lerat (2002a, b) argues for the application of object classes to terminology. The Generative Lexicon is applied by Aldestein and Cabré (2002) to analyze polysemy. Distributional Semantics was applied by Bernier-Colborne (2016) to identify relations between terms and discover sets of potentially frame evoking terms.

CHAPTER 4

What is a term?

It was stated at the beginning of Chapter 2 that terminology is a field that studies terms. Paradoxically (or not, as will be seen further on), there is no real consensus on the notion of 'term', and making a list of relevant terms in a given domain is not an easy task. This difficulty was stressed by many authors, for instance:

> There is no fully operational definition of terms. (Gaussier 2001: 168)

> … the difficulty in distinguishing between two types of units stems from the fact that formally terms are indistinguishable from words. (Sager 1998: 41)

This chapter first examines issues raised by the definition of 'term' and then presents some criteria based on lexical semantics that can help terminologists select which units have terminological status in specialized texts.

4.1 Identification of terms

A short text on endangered species will serve to illustrate some of the questions that arise when trying to identify terms. This extract is no more difficult than any other and the different problems discussed below would inevitably occur in any specialized text.

Typical questions are as follows:

- *Peregrine falcon, habitat* and *extinction* are certainly terms linked to endangered species, but should *population*, *risk* and *river* also make the cut? This first question raises the issue of establishing a relationship between ***lexical items*** or ***linguistic expressions*** and a special ***subject field***.
- A terminologist might include *habitat* in a dictionary on endangered species, but would an expert or a translator consider it to be a term as well? This raises the question of the ***application*** for which terms are collected and described.
- Are *Species at Risk Act* and *Vancouver Island* terms? In other words, can ***proper nouns*** be regarded as relevant terms?
- Are *survive* and *endangered*, which appear in the following sentences, valid terms?

Endangered Animals
Many animals in Canada face the risk of extinction. The major factors that put Canadian species at risk are the destruction of natural habitats, excessive commercial harvesting (e.g. hunting and fishing) and pollution. The most significant challenges for animals are decreases in the amounts of natural forest and grassland as those ecosystems are converted for agricultural and urban uses, as well as commercial timber harvesting and pollution of surface waters (lakes and rivers) by organic matter, nutrients and toxic chemicals. As of 2013, a total of 676 species were considered to be at risk in Canada, including 456 kinds of animals. (Other species at risk include plants.)

Designations
In 2002, the Government of Canada passed the Species at Risk Act, its first endangered species act. Under that legislation, species-at-risk are designated by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC). COSEWIC is funded by Environment Canada, but it otherwise operates independently of the government. COSEWIC commissions studies of native species whose survival in Canada might be at risk, and based on that research it designates them as being in one of several categories of risk: extinct (not surviving anywhere), extirpated (no longer in Canada but surviving elsewhere, usually in the U.S.), endangered (at risk of becoming extirpated or extinct), threatened (at risk of becoming endangered), or special concern (at risk of becoming threatened).

Canadian Case Studies
Sea Otter
The Sea Otter was heavily exploited on the Pacific coast by hunters who sold as many as 1,200 pelts per year during the late 1700s and 1800s. By 1900, sea otters were on the verge of extinction; the last documented sighting in British Columbia was in 1929. An international treaty (1911) gave protection to the endangered sea otters, and by the late 1960s the Alaskan population had grown to about 30,000. Transplants to the west coast of Vancouver Island(1969–72) were successful and the BC population is now several thousand animals. It was downlisted from endangered to threatened in 1996, and in 2007 was further downlisted to special concern.

Peregrine Falcon
Many populations of peregrine falcons were decimated by organochlorine chemicals such as DDT and PCBs, which cause various problems, such as reducing calcium in eggshells so that they break under the weight of an incubating parent. Organochlorines and other human-caused influences resulted in the peregrine falcon population plummeting throughout Canada. However, since the manufacturing and use of organochlorines was banned in the 1970s, populations of this species are now increasing. In 2002, the status of the peregrine falcon was changed from threatened to special concern.

(Freedman & McAllister 2015)

> *… extinct (not* SURVIVING *anywhere), extirpated (no longer in Canada but* SURVIVING *elsewhere, usually in the U.S.).* *(Freedman & McAllister 2015)*

> *Endangered animals*
> *It was downlisted from* ENDANGERED *to threatened in 1996, and in 2007 was further downlisted to special concern.* *(Freedman & McAllister 2015)*

In order words, can ***adjectives*** and ***verbs*** acquire terminological status the same way nouns do? This raises the issue of which ***parts of speech*** can be defined as terms.

- *Survival* and *survive* appear to be denoting the same concept. Should they both be taken into consideration? Do they have the same status? This raises the problem of dealing with ***variants***.
- Are the ***multiword expressions*** *endangered animal*, *natural habitat* and *peregrine falcon* relevant terms, or should one consider *endangered*, *animal*, *natural*, *habitat, peregrine,* and *falcon* separately?
  The answer to this question has consequences with respect to other expressions in the text and other texts on the same subject. If *endangered animal* corresponds to a term; then *endangered species* and *endangered plant* should be defined as terms as well. The same reasoning applies to *natural habitat*: if the entire expression is defined as a term, then *natural forest* must also be regarded as a ***multiword term***.
  In contrast, if *endangered* and *animal* are separate terms; then *species* would probably qualify as a ***single-word term*** as well. The same applies to *forest* and *habitat*. In this case, *endangered* and *natural* when used alone would probably be defined as terms (if adjectives can have terminological status which brings us back to a previous question).
  Peregrine falcon poses a different problem. *Peregrine* does not modify other terms that denote animals in the field. Furthermore, *peregrine falcon* designates a species with specific characteristics: i.e. 'a falcon having gray and white plumage found worldwide and used in falconry'. The meaning of *peregrine falcon* cannot be inferred from the separate meanings of *peregrine* 'coming from another country or land' and *falcon* 'bird of prey of the Falconidae family with long, pointed wings, a hooked beak with a toothlike notch on each side of the upper bill'.
  Finally, if *endangered* is selected as a relevant term, then *at risk*, *threatened*, *extinct* and the verb *to endanger* must also be considered since they are closely related.
  The questions listed in the previous paragraphs illustrate three different problems: Can multiword expressions qualify as terms or should only single-word units be considered? If multiword expressions can become terms, what are the boundaries? Can expressions that convey a ***compositional meaning*** (the whole meaning of the expression corresponds to the sum of its parts) be terms?

Each question and some answers provided in terminology are examined more closely in the following sections.

4.1.1 Relationship with a subject field

If asked to identify terms in the extract that appears in Section 4.1, most terminologists (and non-terminologists for that matter) would probably agree that

ecosystem, *extinction*, and *peregrine falcon* are terms. But what about *fishing*, *population, risk* and *human-caused influence*? Are all these terms? Should they be recorded in specialized resources?

The first difficulty, which arises when identifying terms, is spotting the units or expressions that are relevant from the perspective of a special ***subject field***. This implies that some units need to be set apart from other linguistic units that appear in specialized texts. Terminologists are seldom experts of the fields they are asked to describe and must thus acquire knowledge to be able to make decisions about the terminological status of linguistic units.

At first, they might be guided by their intuition about what seems "unusual" in the form or in the meaning of the units.[25] For instance, the unusual sequence *organochlorine chemical* will probably lead everyone to identify it as a term. However, this is rarely enough to guide term selection, since most terms behave exactly like other lexical units. For instance, *population* denotes an important concept in relation to species, but it is not unusual in the way that *organochlorine chemical* is. In fact, as pointed out by Sager in the quote reproduced at the beginning of this chapter, there is nothing in the linguistic or contextual behavior of terms that sets them apart from other units.

Another parameter that needs to be considered when identifying terms is the way the subject field was defined and delimited prior to term selection. Let us go back to the extract on endangered species. If the field is defined as 'endangered species' or as 'classification of endangered animals' and the corpus is assembled accordingly, there will inevitably be effects on term selection. If endangered species is defined as the subject field, the resulting list of terms will be much longer since terminologists might want to list all the different species that are endangered, the designations used to classify the level of risk, different factors that affect the survival of species, and so on. In contrast, terminologists interested in the classification of endangered animals will simply retrieve terms that designate the statuses given to animals that are at risk, and perhaps, the different species that have been classified according to these statuses. Another possible question

25. In some texts, formal clues can help single out terms. Special fonts (italics, bold) can be used to emphasize important concepts. Other clues can be found in textual structure: for instance, placing important concepts in titles or in enumerations. Finally, linguistic markers such as *is defined as* or *means* introduce definitions for things that are likely to be important concepts in the field. While useful, these clues are not entirely reliable since other important concepts can be expressed in text without being introduced by any kind of formal apparatus. Conversely, some items can be italicized or placed in a specific textual environment but not correspond to terms.

can be raised by the extract: Should the list only include phenomena related to species or should it also take into account the scientific apparatus used to observe them?[26]

Stating that a linguistic item is a ***term*** is considering its meaning from the perspective of a special subject field. There is no such thing as a term in essence; a linguistic unit becomes a term relative to the ***subject field*** in which it is considered (in addition to the way this subject field is delimited). This also means that even common linguistic items can become terms in specialized domains. For example, *risk* becomes a term when it labels a specific phenomenon affecting species despite it being a very common English word.

Finally, a linguistic item can also be a relevant term in different fields of knowledge: *peregrine falcon* is a term in the fields of zoology and falconry in addition to endangered species. This means that terms can be listed in terminological resources devoted to different domains. However, a term is susceptible to the perspective given by the domain (we will come back to this in Section 5.1.5).

4.1.2 The importance of the application

Terminology work can be carried out for different reasons, such as dictionary compilation, translation, knowledge modeling, document indexing, and standardization. ***Applications*** have an influence on the list of terms that are deemed relevant and this fact is now recognized by most approaches to terminology.

The role the application plays in term identification was analyzed empirically by Estopà (2001) who asked medical doctors, indexers, terminologists and translators to go through a medical text and identify linguistic units that they considered to be relevant terms. The results revealed differences at various levels. Each category of experts collected a different number of terms (for instance, terminologists identified a higher number of terms than indexers or medical doctors). Furthermore, experts considered linguistic units of different natures. Indexers only identified nouns; translators extracted longer sequences, such as collocations. Assuming that these categories of experts are involved in different applications (medical doctors are concerned with knowledge transfer; indexers

26. Scholars (Phal 1971; Drouin 2007; Tutin 2007; Granger and Paquot 2009; Hatier 2016) have pointed out that the lexicon used in specialized texts can be divided into three different categories: the general lexicon (*be*, *use*, *large*), the terminological lexicon (*greenhouse gas*), and a transdisciplinary lexicon. The latter category contains units that appear in specialized texts but cut across all or several fields of knowledge. *Model* and *scenario* would probably be defined as transdisciplinary lexical units by these authors.

with accounting for the contents of documents; terminologists with resource compilation; translators with transposing a text and its content from one language to another), the differences show that applications have an effect on the notion of 'term'.

Hence, the application is central to term identification. After defining the subject field, it is the second most important aspect to bear in mind. This means that the two most important factors to consider when identifying terms are extralinguistic: the subject field and the application. Let us now turn to more linguistic considerations.

4.1.3 Can proper nouns be terms?

Specialized texts contain different kinds of nouns: common nouns, such as *animal* and *ecosystem*, and proper nouns. Consider the sentences below:

> *In 2002, the Government of Canada passed the Species at Risk Act, its first endangered species act.* *(Freedman & McAllister 2015)*
> *Transplants to the west coast of Vancouver Island (1969–72) were successful and the BC population is now several thousand animals.*
> *(Freedman & McAllister 2015)*

Do *Species Risk Act* and *Vancouver Island* correspond to terms the way *ecosystem* does? Should they be listed in a specialized dictionary?

There is no straightforward answer to this question and solutions vary from one terminological resource to another. Again, this often depends on the applications for which the resources are compiled. In fact, in some fields of knowledge, ***proper nouns*** must be included: in astronomy, the names of celestial bodies, such as planets, would be impossible to ignore.

Regardless of the choice made regarding the inclusion of proper nouns, their linguistic properties differ from those of common nouns. Common nouns denote classes of objects while proper nouns refer to individual entities. For instance, *island* denotes a land that is completely surrounded by water; *Vancouver Island* refers to an island located in Western Canada. Similarly *linguist* denotes a group of experts who study language while *Igor Meľčuk* refers to a single person who dedicated his life to the study of language. Because of this property, proper nouns cannot be defined as common nouns. *Word* is explained below by providing encyclopedic information, whereas *word processor* is defined in a way that helps us understand to which class of objects it applies.

> *word processor: A computer application used to create, edit, and format documents.*
> *Word: Word processor designed by Microsoft.*

This difference in denotation – classes or individual objects – also results in different linguistic behaviors:

- In English and other languages, many proper nouns cannot be used with a determiner: *a linguist*, **a Eugen Wüster*; an *operating system*; **a Windows*. Others proper nouns accept the definite article, but not the indefinite one: *the Vancouver Island*, **a Vancouver Island*.
- Proper nouns share a very limited set of lexical relations with other linguistic items. For instance, they do not combine naturally with verb or adjective collocates: *a robust operating system; ?a robust Windows*.
- In some languages, like English or French, proper nouns can be distinguished from common nouns because they begin with a capital letter.

Proper nouns include names of people (e.g. *John*, *Wüster*), locations (*Africa, Kyoto*), and trademarks (*Toyota*, *Windows*). Some units that used to function as proper nouns have become common nouns (*a kleenex*, *a PC*) and hence no longer display all or part of the linguistic properties listed above. Most of what is said in this book applies to common nouns.

4.1.4 Different parts of speech

There is a larger consensus around ***nouns*** than around any other parts of speech as far as defining terminological status is concerned. It seems that establishing a relationship with a special subject field is more naturally applicable to nouns, especially nouns that denote entities. For some reason, it is much more difficult to determine the terminological status of verbs and adjectives, such as *survive* and *endangered* in the following sentences.

> *COSEWIC commissions studies of native species whose* SURVIVAL *in Canada might be at risk.* *(Freedman & McAllister 2015)*
> *… extinct (not* SURVIVING *anywhere), extirpated (no longer in Canada but* SURVIVING *elsewhere, usually in the U.S.).* *(Freedman & McAllister 2015)*

> *ENDANGERED animals*
> *It was downlisted from* ENDANGERED *to threatened in 1996, and in 2007 was further downlisted to special concern.* *(Freedman & McAllister 2015)*

Even if a terminologist decides to select *survival* for inclusion in a dictionary of endangered species; *survive* might not be retrieved. The adjective *endangered* might be selected as such, but it will likely be considered within a longer sequence, such *as endangered species*.

This preference for nouns is confirmed in most specialized dictionaries and other terminological resources. Table 4.1 shows how parts of speech are

distributed among entries in a law dictionary called *Legal Dictionary*.[27] A law dictionary is more likely to contain terms that refer to concepts other than entities and this should have an effect on the proportion of verbs and adjectives in the word list. Table 4.1 shows that nouns still largely outnumber other parts of speech (84.3%). Verbs and adjectives represent around 15% of the entries. In other dictionaries, the proportion of nouns can be even higher. In a previous analysis, entries from four specialized dictionaries in different fields of knowledge were analyzed (L'Homme 2003). Nouns or noun phrases counted for 84.25% to 97.5% of the entries. Adjectives came in second (in a medical dictionary, adjectives made up 11% of the sample), and verbs came in third. In one of the resources, not a single verb was recorded.

A closer look at the entries of existing resources reveals that this focus on nouns leads to inconsistencies. For instance, Foldoc (2015) (an online dictionary of computing) contains entries for the noun *program*, for which incidentally a single meaning is recorded, and contains entries for *programmer* and *programming*. However, there is no additional entry for the verb *program*, or the adjective *programmable*, although the latter appears in a multiword unit.

Table 4.1 Parts of speech in the Legal Dictionary (2015)

Noun		**Verb**		**Adjective**		**Adverb**
306 (84.3%)		35 (9.64%)		19 (5.23%)		3 (0.83)
N	NP	V	VP	Adj	AP	
140	166	31	4	18	1	
TOTAL						363

* Different meanings with the same form were counted separately. Cross-references were removed.
One preposition (*contra*, defined as an adjective in the dictionary) does not appear in the table. One abbreviation (C.I.F.) was also omitted.
Two forms (*capricious* (with two meanings) and *careless*) are indicated as both adjectives and adverbs in the dictionary.

These inconsistencies can hardly be justified. How can a unit that is semantically related to another one that is already recorded in a dictionary not be considered? Furthermore, many verbs and adjectives such as *anthropogenic* and *abiotic* (in the field of the environment), and *download* (in computing) can only be defined with reference to a specialized subject.

Deciding to consider nouns only or different parts of speech may depend on the application for which terms are collected. Nevertheless, it seems that

27. The entries under the letter C were analyzed.

most terminological resources do include verbs or adjectives and some adverbs. Although, terms may belong to different ***parts of speech***, they fall into the category of ***open class units***, which have the following properties (according to Mel'čuk & Milićević 2014):[28]

- Open class units are created on a regular basis and some disappear from the lexicon of a language.
- They convey a lexical meaning, i.e. they denote entities, activities, relations, properties; concepts that correspond to the speaker's construal of the real world.
- They include nouns, verbs (other than auxiliaries), adjectives and adverbs.

Hence, terms can be ***nouns*** (*biodiversity*, *temperature*, *computer*), ***verbs*** (*contaminate, download*, *erode*), ***adjectives*** (*extreme*, *green*, *virtual*) and ***adverbs*** (*dynamically*, *globally, online*). However, although the proportion may vary from one terminological resource to another, nouns usually prevail. In addition, only a few adverbs can be considered terms and most are derived from adjectives (*globally*) or have an adjective counterpart (*ONLINE application; go ONLINE*).

Terms may denote ***entities*** expressed typically with nouns, such as *bicycle*, *computer*, *programmer*. In certain constructions, entities can be expressed by a specific class of adjectives, called ***relational adjectives***, such as *viral* in *viral attack*; or *anthropogenic* in *anthropogenic climate change*. Terms can also denote ***activities***, ***events*** or ***processes*** that are expressed by verbs or nouns, such as *cool*, *erode*, *tornado* or *warming*, and ***properties***, expressed by adjectives, nouns or adverbs, such as *clean*, *dynamic*, *green*, *compatibility*, *temperature*.

4.1.5 Single-word items versus multiword expressions

When selecting terms, ***multiword expressions***, such as *endangered animal*, *natural habitat*, are often regarded as the most likely candidates for terminological status. Again, this is consistent with what is done in specialized resources. A quick look at specialized dictionaries reveals that most entries correspond to multiword terms. In a cycling dictionary (Office québécois de la langue française 2018), 77% of the 503 French terms recorded in the word list are multiword terms (e.g. *vélo*

---

28. In contrast, ***closed class units*** have the following features (again according to Mel'čuk & Milićević 2014):

- The creation of new closed class units is uncommon.
- They convey a grammatical meaning; this meaning is very poor semantically.
- They include auxiliary verbs, determiners, prepositions, conjunctions, pronouns, particles.

électrique (electric bicycle), *réseau cyclable urbain* (urban bicycle network)), while the remaining 27% are single-word terms (e.g. *vélo* (bicycle), *roue* (*wheel)*).[29]

Knowledge-based approaches justify the inclusion of multiword terms by the fact that they label items of knowledge (a concept) in a conceptual structure. Hence, it only seems natural that a dictionary on cycling records terms that denote different kinds of bicycles since they all designate specific concepts.

vélo (bicycle)
- vélo (according to its shape)
  - bicyclette pliante (folding bike)
  - monocycle (monocycle)
  - vélo à position allongée (recumbent bicycle)
  - etc.
- vélo (according to its use)
  - vélo pour enfant (children's bicycle)
  - vélo de cyclotourisme (touring bicycle)
  - vélo d'intérieur (exercise bicycle)
  - vélo de montagne (mountain bicycle)
  - etc.
- vélo (according to the way it is propelled)
  - vélo électrique (electric bicycle)
  - vélo hybride (hybrid bike)

From the point of view of lexical semantics, ***multiword terms*** raise a number of issues:

- Many multiword terms are completely compositional. For instance, the expression *endangered species* can be understood by combining the separate meanings of *endangered* and *species* (a species that is endangered). Occurrences *endangered* in a corpus of environment texts show that it can modify other nouns, such as *endangered animal*, *endangered ecosystem, endangered habitat, endangered mammal*, etc. Similarly, *species* can be used alone or in other sequences: *desert species*, *ocean species*, *terrestrial species, threatened species*, etc. In both cases, the meanings remain unchanged. This also applies to the *bicycle* examples given above. A *vélo pour enfant* (children's bicycle) is a bicycle for children, a *bicyclette pliante* (folding bicycle), a bicycle that can be folded, a *vélo électrique* (*electric* bicycle), a bicycle propelled by electricity and so on.
- Assuming that compositional multiword sequences can be terms, how can we delimit them in running text? For instance, in the sequence *endangered ocean*

29. In this count, acronyms (e.g. VTT) were ignored, each meaning of polysemous items were defined as separate terms (*remorque*$_1$, *remorque*$_2$) and compounds were counted as single-word terms (e.g. *porte-patin*).

species, do we have one or multiple terms? Possible readings are: *endangered ocean species, ocean species,* and *endangered species* (in addition to each unit being considered separately).

– Making decisions about which multiword terms should be included in resources may lead to inconsistencies. If *endangered species* is added to an environment dictionary, then *endangered animal* should be listed as a synonym or a variant. *Threatened species* should also be considered. Similarly, if types of bicycles defined according to their use are recorded in a dictionary on cycling, then all must be added (*mountain bicycle*, *city bicycle*, *exercise bicycle*, etc.). This can result in very long lists and terminological resources might leave out multiword sequences at one point or the other.

The solutions to these problems differ quite dramatically whether one takes a knowledge-based or a lexicon-based approach. A knowledge-based approach takes the knowledge structure into account regardless of matters regarding the compositionality of linguistic expressions. In fact, knowledge-driven approaches prefer transparent labels and, because of this, favor compositional multiword terms.

A methodological solution based on lexical semantics will be offered in Section 4.3. It consists in defining terms as lexical units and thus considers that multiword expressions only correspond to terms if their meaning is non-compositional.

It should also be pointed out that some terminological repositories include ***derivational morphemes*** in their word list. This is done when the ***morpheme*** is productive in a domain and conveys a regular meaning. For instance, *bio-* could be listed in an environment dictionary, since it appears in *biodiesel*, *biofuel*, *biogas*, etc. and consistently carries the same meaning. Even if these kinds of entries are useful, we will see in Section 4.3 that *bio-* is not a lexical unit.

4.1.6 Different names for the same thing

Different linguistic means can be used to express the same concept or meaning in running texts, as shown in the following sentences:

> *… most of global warming was being caused by increasing concentrations of greenhouse gases produced by* HUMAN *activities (*ANTHROPOGENIC*).*
> *(Intergovernmental Panel on Climate Change (IPCC) 2006d)*
>
> *Global warming and climate change both refer to the observed century-scale rise in the average temperature of the Earth's climate system and its related effects. (Intergovernmental Panel on Climate Change (IPCC) 2006b)*
>
> *Multiple lines of scientific evidence show that the climate system is warming. (Intergovernmental Panel on Climate Change (IPCC) 2006a)*
>
> *Some causes for change, whether natural or human, tend to cool climate, while others induce warming. (Environment Canada 2007)*

A positive radiative forcing (such as that due to increasing GHG concentrations) WARMS *the atmosphere.* *(Environment Canada 2007)*
**WARMING* *is expected to be strongest in the Arctic, with the continuing retreat of glaciers, permafrost and sea ice.* *(Environment Canada 2007)*

In the first sentence *human* is paraphrased by a more scientific adjective, i.e. *anthropogenic*. In the second series of sentences, *global warming* is replaced alternatively by the noun *warming*, and by the verb *warm*. These expressions have been called ***term variants*** and the phenomenon whereby the names of concepts change is labeled ***terminological variation***.

For terminologists, variants pose two different kinds of difficulties. The first one is being able to recognize linguistic expressions that denote the same concept in running text. Another difficulty lies in knowing which variants should be considered and for what purpose. Some ***variants***, called ***denominative*** (Freixa 2006), are candidates for inclusion in terminological resources. In this sense, variants could be confused with synonyms; however, variants differ from ***synonyms*** since they occur in a variety of forms that include but are not limited to synonyms. For instance, *human* and *anthropogenic* in the examples above could be considered to be denominative variants. Other kinds of ***variants (contextual)*** can help acquire knowledge about a meaning or a concept, but are normally not included in terminological resources as such. The above variants of *global warming* are contextual. The verb *warm* would not be recorded as a valid synonym for *global warming* since it belongs to a different part of speech.

4.2 Different approaches to the 'term'

It should be obvious by now that ***term*** is a relative notion, since different parameters must be taken into consideration, the most important ones being extralinguistic: the subject field and the way it is delimited, along with the application for which terms are collected. This explains why experts produce different lists when asked to retrieve terms from a specialized text. In addition to those factors, various theoretical approaches to the notion of 'term' coexist, and this can also lead to considering different kinds of linguistic expressions.

Over the past decades, in addition to the perspective of the General Theory of Terminology (GTT), for which a term is a label for a clearly defined concept, other views have emerged to characterize the objects terminologists should consider. A short chronological list is given below.

– The term, a unit that varies: ***Socioterminology*** (Boulanger 1995; Gaudin 1993, 2003) was probably the first approach to question the principles of the GTT,

and traces of its influence can now be found in many other approaches. Socio-terminology has drawn the attention of terminologists to the fact that terms are not fixed units and that they vary according to sociolinguistic factors.

– A ***construct***: In an article published in 1999, Bourigault and Slodzian presented a new approach to terminology. In ***Textual terminology***, the 'term' is defined as a 'construct'. Among others, the authors emphasized the role played by the application.

> Le terme est un "construit", c'est-à-dire qu'il résulte de l'analyse faite par le terminographe : cette analyse prend en compte la place occupée par le terme dans un corpus, une validation par des experts et les objectifs visés par une description terminographique donnée. (Bourigault & Slodzian 1999: 31)[30]

– A 'polyhedron' or a 'unit of specialized knowledge': It is now agreed that terms can be viewed from different perspectives triggered by specific applications. Cabré (2003), the main proponent of an approach called the ***Communicative theory of terminology***, accounts for this multiplicity of perspectives by means of a metaphor: the ***polyhedron***. Cabré defines the term as an object that can be envisaged from the points of view of cognition, linguistics, or communication, and that can be studied in frameworks derived from these perspectives. For Cabré (2003), various linguistic and non-linguistic forms (multiword terms, single-word terms, collocations, symbols, etc.) can convey knowledge and correspond to what the author calls ***units of specialized knowledge***.

– A unit grounded in culture: Diki-Kidiri (2000; 2007) challenges the General Theory of Terminology from a different angle. Based on surveys carried out in African countries, the author suggests that languages – and consequently terms that are part of these languages – are influenced by the cultures in which they are grounded. Hence, when revitalizing languages, cultural factors should be taken into account. The principles advocated by Diki-Kidiri led to the ***Cultural approach***.

4.3 Terms as lexical units

Although intuitively attractive, the approaches to the notion of 'term' listed in the previous section only help us indirectly manage the difficulties mentioned earlier

---

30. "The term is a "construct", i.e. it results from the analysis carried out by a terminologist: this analysis takes into account the place of the term in a corpus, a validation by an expert, and the objectives of a specific terminological description". (my translation)

in this chapter. More specifically, they do not provide a formal basis to distinguish terms from other lexical units, and more precise and operational criteria are needed. We must also ensure that the criteria can be applied consistently to all linguistic expressions that occur in specialized texts.

Our approach to ***terms*** is to consider them as ***lexical units*** (LUs) as defined in lexical semantics. This approach will be applied systematically in this book when referring to lexicon-driven approaches. This section presents well-known tests used by linguists to differentiate lexical units from other linguistic units: from morphemes, on one side, and from phrases, on the other.[31]

Firstly, *lexical unit* is preferred to *word* due to the vagueness of the latter notion. Many linguists have discussed problems raised by the definition of ***word*** that appear both at a formal and at a semantic level. They are illustrated with the following examples:

(1) *I worry when he drives in bad* WEATHER.
This WEATHER *is unpredictable.*
He is under the WEATHER.

(2) *I was standing by the* WINDOW *when Julien called.*
Open this WINDOW *to visualize the icon.*
The cat chases the MOUSE.
Move the MOUSE *over this icon to obtain more information about this option.*

(3) *Reforestation activities are deemed insufficient in order to reforest large areas of Quebec.*
You will need to reinstall this program to have it work properly.

Sentences in (1) illustrate how the delimitation of words can be problematic at a formal level. If only graphical boundaries such as spaces and punctuation signs are considered, the first sentence can be said to contain 8 words, the second one, 4 words, and the third one, 5 words. However, using these basic boundaries produces odd results when applied to the third sentence. It would then be difficult to explain the respective meanings of *under* and *weather* in *under the weather* and their individual contribution to the sentence.

Sentences in (2) show how identical 'words' may convey different meanings. The first and second sentences both contain the word *window*, but each occurrence of *window* refers to a different object: $window_1$ 'an opening in a building or vehicle'; $window_2$ 'a graphical display on a computer screen containing icons and/or files that a user can select'. Similarly, in the third and fourth sentences, *mouse* refers to two separate things: an animal and a peripheral device.

---

31. Readers familiar with these criteria can browse this section quickly. Section 4.4 focuses on criteria to characterize the subset of lexical units that interest this book, i.e. terms.

Finally, sentences in (3) illustrate a formal and semantic problem. *Re-* in *reforestation* and in *reinstall* conveys a meaning of its own, i.e. 'once again'. It can be used in other combinations and still have the same meaning (e.g. *reinforce*, *reprogram*). However, speakers of English would hardly consider *re-* to be a word. How can we distinguish units such as *re-* from words if they convey similar meanings?

Criteria and tests may be applied to overcome these problems and help us identify linguistic expressions called *lexical units*.[32] The criteria given below are based on Cruse (1986). A first set helps us make decisions with regard to the formal level. (In Cruse, criteria applied at this level are said to contribute to the ***syntagmatic delimitation*** of LUs.)

A. A lexical unit should correspond to a ***semantic component***, i.e. a component that conveys a meaning that contributes to the global meaning of a sentence when combined with other components.
Hence, according to this first criterion, *weather* in *I worry when he drives in bad weather* and *this weather is unpredictable* corresponds to a semantic component; however, *weather* in *he is under the weather* does not correspond to a semantic component. *Under the weather* as a whole needs to be defined as a semantic component.
There are different ways to verify the correspondence between unit and semantic component.

   – Test 1: The contribution of a semantic component to the global meaning of a sentence is the same in other sentences.

     I worry when he drives in bad WEATHER.
     This WEATHER *is unpredictable.*
     This area witnesses very unusual WEATHER.

     He is under the WEATHER.

     In the first three sentences, *weather* denotes a specific state that an area experiences at a given time with respect to temperature, dryness, sunshine, wind, etc. *Weather* in *under the weather* does not have the same contribution as in the first three sentences. To be able to observe a similar contribution between items in the last sentence and other ones, we would need to consider *under the weather* as a whole, as in *They have been feeling under the weather for a few days.*

   – Test 2: The replacement of one LU with another gives way to a recurrent semantic contrast. *Conditions* and *temperature* are used to replace

32. From now on, ***word*** will be used to refer to the graphical unit delimited by spaces and/or punctuation marks. .

weather in the following sentences. They introduce meanings that differ from that of *weather*. However, the difference remains the same throughout all sentences.

I worry when he drives in bad WEATHER.
I worry when he drives in bad CONDITIONS.
I worry when he drives in bad TEMPERATURE.
This WEATHER *is unpredictable.*
These CONDITIONS *are unpredictable.*
This TEMPERATURE *is unpredictable.*
This area witnesses very unusual WEATHER.
This area witnesses very unusual CONDITIONS.
This area witnesses very unusual TEMPERATURE.

This test can also be used to validate the fact that *under the weather* as a whole is a lexical unit. We must replace the entire expression to be able to observe and maintain a contrast.

He is under the WEATHER.
He is SICK.
He is ILL.
He is HEALTHY.

They have been feeling under the WEATHER *for a few days.*
They have been feeling SICK *for a few days.*
They have been feeling ILL *for a few days.*
They have been feeling HEALTHY *for a few days.*

B. A lexical unit must correspond to a least one graphical word and this word must have some autonomy in a sentence. In other words, the LU can change positions without affecting the grammaticality of the sentence provided that these positions are compatible with the syntactic valence of the LU. For instance, *reforestation* can appear in different positions in which a noun is allowed:

Reforestation activities are deemed insufficient in order to reforest large areas of Quebec.
In order to reforest large areas of Quebec, REFORESTATION *is deemed insufficient.*
Activities, including REFORESTATION, *are deemed insufficient …*
In order to carry out REFORESTATION *of large areas*

However, according to this criterion, *re-* in *reforestation* is not a lexical unit, since it can only appear in one specific position:

**Forestation* RE- *activities*
**Forestation activities* RE-
Activities of REforestation

Lexical units are also the largest units that accept insertion. Hence, we can modify *reforestation:*

> *Massive* REFORESTATION *of large territories…*
> *Massive* REFORESTATION *is deemed…*
> *Massive and intense* REFORESTATION *of all these territories*

However, this is impossible with *re-* in *reforestation*:

> **Re(aff-)forestation activities*
> *Reforestation (and afforestation) activities are (currently) deemed (largely) insufficient in order to (fully) reforest large areas of (the province of) Quebec.*

A second set of criteria addresses the delimitation of LUs at the semantic level and deals with the problem raised by *mouse* and *window* that are associated with two different meanings. A single criterion is mentioned here since more are presented in Chapter 5. For instance, we can try replacing *mouse* and *window* by lexical units that belong to the same paradigm. If the replacement works in some sentences but not in others, then we have different lexical units.

> *The cat chases the* MOUSE.
> *Move the* MOUSE *over this icon to obtain more information about this option.*
> *The cat chases the rat.*
> *?Move the rat over this icon to obtain more information about this option.*

> *I was standing by the* WINDOW *when Julien called.*
> *Open this* WINDOW *to visualize the icon.*
> *I was standing by the door when Julien called.*
> *?Open this door to visualize the icon.*

Hence, the lexical item *mouse* corresponds to two different lexical units in the sentences above: $mouse_1$ and $mouse_2$. Each has its own definition and shares relations with different sets of lexical units: $mouse_1$ is related to *rat*, *rodent*, *animal*, etc.; $mouse_2$, is linked to *keyboard*, *peripheral*, *click*, etc. Similarly, $window_1$ belongs to the same semantic domain as *door*, glass, *building*, etc.; whereas $window_2$ is linked to *interface*, *file*, *icon*, *screen,* etc.

This operational approach to the notion of 'term', i.e. defining it as a lexical unit according to formal and semantic criteria, has a number of important consequences regarding the way terms are usually envisaged in terminology.

– This ***approach*** is ***semasiological***: A potential terminological meaning attached to a lexical item is discovered based on the analysis of this item as it occurs in text. Hence, the goal is not to unveil the knowledge structure of a subject field, but to understand the specialized meaning(s) that a lexical item may have.

- Considering terms as lexical units is assuming that terms are a part of the ***lexicon*** of a language. Hence, this approach is compatible with most lexicon-driven approaches and leads to a situation in which analyses of general lexical units and terms can be combined and/or unified. However, the traditional goal of terminology that consists in establishing a correspondence between knowledge, its organization and the way it is expressed cannot be met with purely lexico-semantic criteria. Although knowledge is the starting point of any terminological analysis (as was mentioned in Section 3.1.1), decisions on the nature of lexical items are based on lexico-semantic criteria. This allows us to take into consideration LUs that are often overlooked in strictly knowledge-based approaches, such as verbs, adjectives and adverbs, and to take into account their specific linguistic properties.
- The objects considered by knowledge-driven and lexicon-driven approaches differ to a great extent. First, only multiword expressions with non-compositional meanings are considered to be relevant terms in perspectives based on the lexicon. For instance, *operating system* is regarded as a term since we cannot explain its meaning by the sum of its parts. However, *configuration file* or *natural habitat* are not defined as terms since they can be understood based on the meanings of their individual parts.
- This approach allows us to take into account various kinds of relations between terms. However, the relations considered are those that can be observed between the meanings of terms (paradigmatic and syntagmatic relations) and not those that hold between concepts in a knowledge structure.

4.4 Criteria for selecting terms

If terms are lexical units (and, thus, belong to the ***lexicon***), we will need to find out what is special about them and what sets them apart from other lexical units. Here are four criteria that can be applied after determining that a linguistic item corresponds to an LU as seen in the previous section.

The first criterion is knowledge-driven and is the starting point to determine the ***terminological status*** of a lexical unit. The last three criteria are based on a lexicon-driven approach. All criteria can apply to lexical units that belong to different parts of speech, but some are more directly relevant to predicative units.

A. **Link with a special subject field**
   This first criterion is the one usually applied in terminological analysis and requires at least basic knowledge in the field. In other words, it relies on

extralinguistic knowledge. It is more naturally applicable to units that denote ***entities***. For instance, units such as *processor*, *bus*, *chip*, *data*, *keyboard*, *monitor* will not lead to much discussion about their terminological status in the field of computing. Similarly, *bicycle, cyclist, wheel,* and *pedal* are likely to be selected when compiling a cycling dictionary.

However, this first criterion is less easily applicable to other units, such as *type*, *transfer*, *display*, *process* (in computing) or *ride*, *cycle*, or *brake* (in cycling). These verbs are likely to raise much more discussion about their terminological status, even though they express typical activities associated with entities. Principles borrowed from lexical semantics can assist terminologists for these other units. The next three criteria will show how. These criteria can help prevent some inconsistencies that were mentioned earlier in the word lists of specialized dictionaries.

B. **Nature of the arguments**

Given a ***predicative lexical*** or ***quasi-predicative unit***,[33] if its ***arguments*** are realized in the form of terms (identified using criteria A, C or D), the said unit probably corresponds to a term. Consider these sentences extracted from a corpus on climate change in which the verb *warm* shows.

> *A positive **radiative forcing** (such as that due to increasing GHG concentrations) warms the **atmosphere**. (Environment Canada 2007)*
> ***gases** such as carbon dioxide (CO_2) which warm the Earth's **surface** …. (EUROPA 2007)*
> *Sooty **aerosols**, for example, tend to warm regional **climates**. (Environment Canada 2007)*
> *Positive **radiative forcings** tend to warm the Earth's **surface** and lower **atmosphere**. (Intergovernmental Panel on Climate Change (IPCC) 2006b)*
> *the ocean **surface** was warmed by solar **heat** and a gradually increasing **greenhouse effect**. (Environment Canada 2007)*
> ***Energy** from the sun warms the Earth's **surface**. (EUROPA 2007)*

Arguments of *warm* are realized in these sentences with units (*gas*, *greenhouse effect*,[34] *atmosphere*, *climate*, etc.) that would have been identified as terms according to criterion A. Hence, *warm* should be regarded as a possible term in climate change.

The criterion can be used with verbs that are prototypical predicative units, but also with other parts of speech, such as nouns and adjectives. The

33. More will be said about predicative and quasi-predicative units in Chapter 6.

34. *Greenhouse effect* is non-compositional and corresponds to a lexical unit.

following sentences show how the arguments of the noun *impact* and the arguments of the French adjective *compatible* are realized in a corpus on climate change and a corpus of computing respectively.

> *Metrics may include* ***climatic*** IMPACT *on existing agricultural, water, and mineral* ***resources*** *….* *(Schwartz & Randall 2003)*
> *Theories concerning the* ***climatic*** IMPACT *of such* ***emissions*** *vary.* *(Müller & Buchdahl 2000)*
> *It so happens that the* IMPACT *of climate* ***change*** *on world heritage* ***sites*** *will be the subject of a meeting taking place at UNESCO just two months from now, on 16 and 17 March.* *(UNESCO 2006)*

> *… le* ***moniteur*** *doit être* COMPATIBLE *avec la* ***carte*** *graphique … (En. The monitor should be compatible with the graphical board).* *(St-Pierre 1988)*
> *En contrepartie la* ***FAT32*** *n'est pas* COMPATIBLE *avec les* ***versions*** *de Windows antérieures à la version OEM Service Release 2.* *(Initiation au dépannage informatique 2002)*

C. ***Morphological relations***

If the lexical unit is morphologically and semantically related to a lexical unit already identified as a term according to criteria A, B or D, then it is most probably a term. According to this criterion, if *pollution* and *pollutant* are defined as terms in the field of the environment, then *to pollute*, *polluted*, *polluting*, and *polluter* should all be considered since they are morphologically and semantically related to *pollution* and *pollutant*. Similarly, if *editor* is defined as a term in the field of computing, then *to edit* and *editing* should also be considered.

It is also important to emphasize that there must be a semantic relation between the units in addition to a morphological relationship. Some linguistic units can be formally related without sharing a close terminological relationship. For instance, although *application* and *apply* are morphological related, their semantic link is lost in the field of computing since *application* denotes a type of program, not the action of applying.

D. ***Paradigmatic relations***

If the LU is paradigmatically related to a lexical unit that has already been recognized as a term according to criteria A, B or C, then the lexical unit is most probably a term. The criterion aims to cover the paradigmatic relations that are not covered by criterion C, such as synonymy or near-synonymy, antonymy, meronymy, and to a certain extent hypernymy.[35] Examples below show how the criterion is applied to terms that share some of these paradigmatic relations.

35. These relations and others are presented in detail in Chapter 7.

Synonymy or near-synonymy
If *email* is defined as a term in the field of computing, then *message* must be considered as well.
If *Earth* is defined a term in the environment, then *planet* and *globe* must be considered.

Antonymy (or other forms of opposition)
If *open* (as in *open a file*) is considered to be a relevant term in computing, then *close* must be considered.
If *biological* is considered to be a relevant term in the environment, then *chemical* should be a term as well.

Meronymy
If *biome* is defined as an environment term, then *ecosystem* should be considered.
If *species* is considered to be a relevant term in the environment, then *population* and *community* are terms as well.

4.5 Applying term identification criteria to a specific domain

The criteria presented in Section 4.4 can be applied to units in running text, sets of concordances or to lists of units extracted automatically from a specialized corpus. This section examines how they can be used concretely for identifying terms related to an environmental topic, i.e. endangered species. It is assumed that a specialized corpus containing texts similar to the extract reproduced at the beginning of this chapter was compiled and that the criteria are applied to lexical items that appear in this corpus. It is also assumed that lexical items have a certain number of attestations in different texts of this corpus.

The following units would undoubtedly be identified according to Criterion A. (= Link with a special subject field). They denote entities that are directly related to endangered species. These can be used as starting points leading to the identification of other terms.

animal
species
habitat

Criterion B. (= ***Nature of the arguments***) can be applied to some predicative and quasi-predicative units. Looking at instances of *endangered* in a corpus on endangered species, reveals that the realizations of its arguments correspond to terms.

endangered
However, in some situations, these animals prey on livestock, ENDANGERED ***species****, and pets or pose a threat to human health and safety.*
(Wildlife Services 2001)

For ENDANGERED ***birds, reptiles****, and* ***mammals****, the sequential selection of counties on the basis of the unique species they contain leads to a steady increase in the number of populations of each* ENDANGERED ***species*** *already included in the counties sampled (Fig. 3C).* *(Dobson et al. 1997)*

I believe these relatively inaccessible headwater areas and shallow nearshore environments may provide previously unknown foraging and nursery grounds for ENDANGERED ***green turtles****.* *(U.S. Fish & Wildlife Service 2008)*

The terminological statuses of *bird*, *reptile* and *mammal* can be established with Criterion A., but also with Criterion D. (= ***Paradigmatic relations***) since they designate classes of species. Criterion D. can also help us determine the status of *green turtle* since is designates a type of animal. *Green turtle* is non-compositional since it cannot be only defined as 'a turtle that has a green color'. Rather, it is a species defined as 'a large usually herbivorous sea turtle (*Chelonia mydas*) of warm waters with a smooth greenish or olive-colored shell' (Merriam-Webster Online Dictionary 2018).

Criterion B. can also help us determine that *survival* is a likely candidate for terminological status, since its arguments take the form of terms as shown in the examples below. Again, the terminological statuses of some of the realizations of the arguments (*vertebrate*, *invertebrate*) can be validated by other criteria.

Survival

Authority of the Convention on International Trade in Endangered Species (CITES) to assess whether trade or utilisation is likely to be detrimental to the SURVIVAL *of a* ***species****.* *(Endangered Wildlife Trust 2010)*

Gopher tortoise burrows are essential to the SURVIVAL *of a wide variety of* ***vertebrates*** *and* ***invertebrates****, including some that are found nowhere else.* *(Mississippi Museum of Natural Science 2014)*

… determination of the behaviour and SURVIVAL *of* ***animals*** *relative to disease or parasites ….* *(Cooke 2008)*

Then we can use Criterion C. (= ***Morphological relations***) to find other relevant terms linked to the ones we have identified so far: *endangered*, *habitat*, *species* and *survival*. In addition to being morphologically related, these candidates are semantically related.

endangered: danger, endanger
habitat: inhabit, microhabitat
species; subspecies
survival; survive

Finally, criterion D. (= ***Paradigmatic relations***) allows us to identify many more candidates that are linked to the terms we identified so far as shown below.

Synonymy or near-synonymy
endangered: threatened
danger: risk, threat
endanger; threaten
habitat: environment, site, territory

Antonymy or other forms of opposition
animal: plant
survive: thrive

Meronymy
animal: population, community

Hyponymy and hyponymy
animal: green turtle, Peregrine falcon, sea otter, whale
species: bird, fish, invertebrate, mammal, reptile, vertebrate

The statuses of some of these new candidates can be further confirmed by other criteria. For instance, as *threatened*, *survive*, *threat*, *risk* and *threaten* are predicative terms, their terminological status can be further verified with Criterion B. in other words, by examining the realizations of their arguments. Furthermore, these new candidates lead to the discovery of new terms. Hence, terminologists can proceed by applying these criteria to new units until they exhaust the list of possible terms in a domain.

Summary

When identifying terms in running text, terminologists must differentiate them from units that do not have terminological status. This task is not devoid of difficulties since the notion of 'term' is relative to the delimitation of a special subject field and to an application. When asked to identify terms in texts, different experts produce diverging lists.

Nowadays, different approaches to the notion of 'term' coexist. The approach taken here considers terms as lexical units. This entails that they can be delimited syntagmatically and semantically. This perspective has a number of consequences for term selection and contrasts sharply with the perspective taken by most knowledge-driven approaches.

Four criteria can be used to assist the selection of terms: the relationship with a field of knowledge, the nature of arguments, morphological and semantic relatedness and paradigmatic relations. Terms are open-class units and can belong to one of the four following parts of speech: noun, verb, adjective, and adverb.

This being said, establishing the terminological status of a lexical item involves a thorough analysis of its meaning. In some cases, it also involves making meaning distinctions if the lexical item is polysemous. This issue and related ones are the subject of Chapter 5.

Further reading

Taylor (2015) is a compilation of contributions that illustrate numerous perspectives that can be taken on the 'word'.

The notion of 'term' has been discussed by scholars such as Guilbert (1973), Petit (2000), Rey (1976) and Sager (1998). Estopà (2001) is an empirical study that shows how different categories of experts define and characterize terms. L'Homme (2005) reviews the various theoretical viewpoints on the notion of 'term' and then presents a lexicon-driven approach.

The ideas of Socioterminology are presented in Gaudin (1993, 2003). Unfortunately there are very few publications that describe the general ideas of Textual terminology as such (refer to Bourigault & Slodzian 1999). To find more about the Communicative theory of terminology, read Cabré (2003). Finally, Diki-Kidiri (2000) presents the general principles of Cultural terminology.

One of the first attempts to isolate a 'general scientific lexicon' in French was carried out by Phal (1971). Other studies were conducted by Drouin (2007), Granger and Paquot (2009), Hatier (2016), and Tutin (2007). In English, a well-known list of academic vocabulary was published by Coxhead (2000).

The first alternative terminological framework to focus on variation was socioterminology (Boulanger 1995; Gaudin 1993, 2003). Since then, terminological variation has been the topic of a wealth of publications. Typologies of variants can be found in Daille et al. (1996) and Daille (2017). Causes of terminological variation are the focus of Freixa (2006). Different viewpoints on terminological variation are offered in Drouin et al. (2017).

Verbs are the focus of an increasing number of terminological studies (L'Homme 1998; Lerat 2002b; Lorente 2007; Pimentel 2013; Ghazzawi 2016). Adjectives in specialized corpora have attracted less attention; still some interesting analyses can be found (Carrière 2006; Maniez 2015; Pitkänen-Heikkilä 2015).

CHAPTER 5

Concepts, meaning and polysemy

In Chapter 3, we saw that different perspectives can be taken on ***linguistic content*** and that they have a direct impact on the way terms are defined, linked to one another and described in resources. This chapter explores the matter further by using a distinction made earlier between knowledge-based and lexicon-based approaches to terminology. Of course, things are not so clear-cut in practice, but this pedagogical distinction allows us to better highlight the consequences of taking one approach or the other when considering linguistic content. It also allows us to understand why each one offers different explanations for similar phenomena.

5.1 Knowledge-based approaches to linguistic content

Previous chapters have explained that the perspective on meaning of the General Theory of Terminology (GTT) is firmly grounded in knowledge as it is defined by experts in special subject fields. This view is shared to a certain extent by more recent knowledge-based approaches to terminology. Let us recall some of the GTT's principles in a way that will be useful for the material presented in this chapter:

- The focal point of analysis is the ***concept***; a concept corresponds to an ***item of knowledge*** and its definition relies on a consensus reached at a certain point in time.
- A concept is assumed to be a generalization of objects in the real world, provided, of course, that objects of concern are linked to a specialized domain;[36]

36. As was mentioned in Chapter 3, many knowledge-based approaches to terminology have forged their own definition of 'concept' that is still prevalent in many circles, but we will see later in this chapter that other alternative views are borrowed from cognitive linguistics to explain the nature of the 'concept'.

- The concept is approached in terms of ***necessary and sufficient conditions*** which correspond to ***characteristics*** that experts deem relevant to define the concept.
- The concept is the starting point of the analysis according to an ***onomasiological approach***; designations are linked to this concept once it is delineated.

These principles allow terminologists to classify different concepts in clear-cut and non-overlapping categories. Once defined, a concept appears in a ***conceptual structure*** in which it is distinguished from other concepts and with which it holds different types of relations. The perspective has important theoretical and methodological consequences that are presented in the following sections.

5.1.1 Dealing with multiple concepts

The GTT advocates ***univocity*** since it seeks to contribute to unambiguous specialized communication. This principle implies that there should be a one-to-one mapping between concepts and designations: a concept should have a single designation and, conversely, a designation should correspond to one concept. In recent years, a plethora of studies have shown that variation is an essential part of specialized communication and argue that it should be managed rather than controlled. Other studies have shown that the reverse situation – a designation can denote more than one concept – can also occur in specialized domains. Hence, the relationship between concepts and designations is many-to-many (rather than one-to-one) as shown in Figure 5.1.

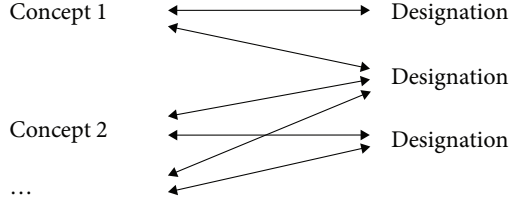


Figure 5.1 Possible relationships between concepts and designations

This being said, it is quite reasonable to assume that the number of cases where a designation can label multiple concepts is reduced when looking at matters from the point of view of specialized fields. However, these cases do occur as explained below.

First, the same designation can label different concepts in distinct fields of knowledge. For instance, in medicine, *virus* denotes 'an infectious agent that replicates only within the cells of living hosts' (*a virus which can kill humans*). In

computing, *virus* is defined as 'piece of code that replicates by inserting copies of itself into computer programs, files, or hardware components' (*computer VIRUSES are malicious programs*). Similarly, *environment* can be defined as 'the set of physical, chemical, and biotic factors (as climate, soil, and living things) that act upon an organism or an ecological community and ultimately determine its form and survival' (Merriam-Webster Online Dictionary 2015) (*impact on the ENVIRONMENT*). In computing, it can refer to 'the entire set of conditions under which a user runs a computer, as it relates to the hardware, operating platform, or operating system' (*a UNIX ENVIRONMENT*).

Traditionally, terminology has considered this first situation as one of ***homonymy***.[37] ***Polysemy*** would occur when a designation labels different concepts in the same subject field as illustrated below with *gas*.

A second situation occurs when the same designation is associated with different concepts in the same subject field. For instance, in the environment, *gas* can denote 'a state of matter that has neither shape nor volume that accumulates in the atmosphere' (*These gases spread all over the Earth*), but also 'a substance used to produce energy' (*In general, the more energy the battery is capable of delivering, the greater the GAS fuel savings*). As was said above, the GTT advises against polysemy. In this case, one option would be to select *gasoline* as the preferred term for the second concept.

Of course, there are also cases in which interferences with general language occur. A lexical item can denote a concept in a specialized field and convey a different meaning in everyday situations. The item *mouse*, for instance, would most readily be associated with the following meaning 'a small rodent' (*there is a MOUSE in the house*) in general language. In computing, *mouse* is used to designate 'a peripheral device used to interact with the computer' (*select this icon with your MOUSE*). The last situation is of interest to terminologists only to the extent that attempts are made to explain connections between general language and specialized languages. This will be dealt with in Section 5.2.1.

5.1.2 Accounting for concepts in terminological resources

Standard terminological resources, such as term banks, deal with homonymy and polysemy in a straightforward way. Section 2.2 already explained that each term record in a term bank accounts for a single well-defined concept. This implies that when the same designation is used to label different concepts in the same or

37. We will see later that homonymy is defined differently in lexical semantics.

in different fields of knowledge, it should appear in separate term records.[38] For example, the Canadian term bank TERMIUM Plus® (2017) has different records for the term *virus*. Among these, one is associated with the fields of microbiology and parasitology; another is associated with *computer programs and programming* (Figure 5.2). In other cases, terms that pertain to different domains can be grouped on the same record. For instance, *carbon dioxide* is described in a single term record in TERMIUM Plus® (2017) but it is linked to the domains of Chemical Elements and Compounds, Biochemistry, Pollutants, and Atmospheric Physics.

Computer programs and programming	**Microbiology and parasitology**
computer virus, ***virus****, electronic virus*	*Virus*
Definition. A program that propagates itself by modifying other programs to include a possibly changed copy of itself and that is executed when the infected program is invoked.	Definition. [A] minute infectious [agent] characterized by a lack of independent metabolism and by the ability to replicate only within living host cells.

Figure 5.2 *Virus* associated with different domains in TERMIUM Plus® (2017)

We saw that the same designation can be used to label different concepts in the same specialized field. Again, the concepts should be described separately. For instance, *Le grand dictionnaire terminologique* (2015) records the French noun *connexion* in eight different term records associated with computer science. Among these records, one describes an activity concept while another defines *connexion* as a set of hardware and software components that allow communication between components of a computer system.

In specialized dictionaries, multiple meanings are often presented in the same entry instead of separate entries. However, meanings are usually listed with no attempt to highlight potential connections between them. As shown below, two meanings of *key* in computing are described in Webopedia (2015), a specialized resource dedicated to computing. Similarly, *acceptance* is associated with four definitions in the *Legal Dictionary* (2015).

> **key**
>
> 1. A button on a keyboard.
> 2. In database management systems, a key is a field that you use to sort data.
>
> (Webopedia 2015)

38. This also implies that if multiple designations exist for a concept, they are all listed on the same term record. This is shown in Figure 5.2 where *computer virus,* ***virus****, electronic virus*, used to denote 'a type of malicious program', all appear on the same term record. It was also shown in Figure 2.4 for the concept 'global warming'.

> **Acceptance** n.
>
> **(1)** receiving something from another with the intent to keep it, and showing that this was based on a previous agreement. **(2)** agreeing verbally or in writing to the terms of a contract, which is one of the requirements to show there was a contract (an offer and an acceptance of that offer). A written offer can be accepted only in writing. **(3)** receiving goods with the intention of paying for them if a sale has been agreed to. **(4)** agreement to pay a bill of exchange, which can be an "absolute acceptance" (to pay as the bill is written) or "conditional acceptance" (to pay only when some condition actually occurs such as the shipment or delivery of certain goods). "Acceptance" is most often used in the factual determination of whether a contract was entered into. (Legal Dictionary 2015)

The distinction by domains made by term banks sometimes leads to a situation where the same concept is described in more than one term record. In TERMIUM Plus®, there are more than 30 entries for *key*. (Table 5.1 shows part of the domains with which the designation is associated.) It would be difficult to argue that *key* actually corresponds to 30 different concepts. This is partly a reflection of the different phenomena presented in this chapter, such as homonymy, polysemy and multidimensionality, and partly the result of concerns having to do with data management.

5.1.3 Explaining concepts

Up to now, we have examined how concepts are considered in knowledge-based approaches and how they are dealt with in terminological resources, such as term banks. Once a specific concept has been identified and delimited, how should it be explained?

Most knowledge-based approaches consider meaning in terms of ***necessary and sufficient conditions***. According to this view, a well-formed ***definition*** should be formulated to include all the members of a class and only those members. In other words, a good definition for 'mouse' in computing should be valid for a broad range of objects that fall into this category (optical mouse, mechanical mouse, wireless mouse, small mouse, large mouse, etc.), but should not include other input devices, such as the keyboard or the touch pad. The definition should list ***characteristics*** such as 'peripheral', 'used to enter data into the computer', 'that can be moved on a flat surface', and 'that is shaped in such a way that the user can move it with his hand'. 'Peripheral' includes devices such as the mouse, but also the printer and the monitor; 'Used to enter data' includes the mouse, but also *keyboard*, 'that can be moved on a flat surface' and 'that is shaped in such a way that the user can move it with his hand' applies to mouse and the different shapes in which it comes but no longer to the keyboard. Looking at matters from the point of view of traditional terminology, this method seems like a sensible option,

Table 5.1 *Key* in TERMIUM Plus® (2017): partial list

Items in the entry	Part of speech	Domain(s)	Definition
key	Noun	Telephone Switching, Intercoms, Telephones	A hand-operated switching device ordinarily comprising concealed spring contacts …
key	Noun	Machine Shafts, Journals and Swivels; Joining Elements (Mechanical Components); Aircraft Systems	A metal piece shaped like a prism – either exactly so or with a slight taper …
key	Noun	Compartment – ISO/IEC JTC 1 Information Technology Vocabulary	<computer security> bit string that controls the operations of encryption or decryption.
key	Noun	Compartment – ISO/IEC JTC 1 Information Technology Vocabulary	<organization of data> identifier that is part of a set of data elements
key, key in	Verb	Information Processing (Informatics)	To enter data into a computer system using any device with a keyboard.
key	Noun	IT security	Information used to set up and periodically change the operations performed in crypto-equipment …
legend, key	Noun	Statistical Graphs and diagrams	An explanatory list of the symbols on a … chart.
key	Noun	Telecommunications, Electrical Appliances and Equipment, Telegraphy, Telecommunications Transmission	A specialized hand-operated switch used to make and break a circuit.
key	Noun	Wind instrument	Metal lever on wind instruments that opens or closes air holes and varies the pitch.
…			

since concepts are clearly defined, do not overlap, and are distinguished from one another with carefully chosen characteristics.

Defining a ***conceptual structure*** with methods, similar to those described briefly in Chapter 3, helps terminologists identify relevant characteristics to include in definitions. The definition should reflect the position held by concepts in the conceptual system. The first defining element should correspond to the generic concept, as shown in the examples below. The other parts of the definition should inform us on how the concept differs from its neighbors.

Mouse: 'Input device' operated by moving it on a flat surface used to control the cursor and select icons and files on a computer screen.
Keyboard: 'Input device' that consists of a series of keys used to enter data into a computer.
Wireless mouse: 'Mouse' with no cord that transmits data to a computer via infrared radiation.
And so on.

Although different types of definitions can be found in specialized resources (Sager 1990), the preferred one is the ***analytical definition***. This definition fills the conditions that were just mentioned. It comprises a ***genus*** – the part of the definition that corresponds to the superordinate concept – and ***differentia*** – the remainder of the definition that specifies how the concept differs from other concepts that are linked to the same superordinate concept. In the 'wireless mouse' example above, 'mouse' is the genus; 'with no cord' and 'that transmits data to a computer via infrared radiation' are the differentia.

5.1.4 An alternative view on concepts

The discussion about specialized concepts and terminological definitions presented so far in this chapter hinges on the assumption that the concepts can be explained with a finite list of characteristics (or, in other words, necessary and sufficient conditions). This list exhausts their contents, distinguishes them from one another and assigns them to a precise position in a conceptual structure. These are very strong assumptions that can raise difficulties in practice when writing definitions for a wide variety of different concepts. This section explains why.

Making a list of ***necessary and sufficient conditions*** for all the members of a class is quite a difficult task and perhaps even an impossible one for certain concepts. One spectacular example was given by Wierbicka (1985): if we had to define the concept 'bicycle' in terms of necessary and sufficient conditions, what would these conditions be?[39]

- 'A small vehicle': This includes bicycles, but also many other vehicles, such as small motorbikes, tricycles, scooters.
- 'Propelled by human energy': This includes bicycles, tricycles, scooters, but no longer motorbikes; it does not include electric bicycles either.
- 'Has two wheels': This includes bicycles and some scooters, but no longer tricycles; it also sets aside stationary exercise bicycles.

39. What follows is an adaptation of the original arguments given by the author.

– 'Has two pedals': This includes bicycles, tricycles, but rules out scooters.
– 'Has a saddle seat': This includes standard bicycles, but vehicles on which cyclists lie down horizontally are excluded.

This short list shows that adding conditions excludes objects that would probably qualify as bicycles: for instance, the 'saddle seat' or 'two wheels' conditions rule out small vehicles on which users lie down and exercise bicycles. However, if we remove some conditions, we might include vehicles that would not normally be regarded as bicycles. Furthermore, this view does not take into account the fact that a bicycle lacking a wheel or a pedal would still be recognized as a bicycle by most people. Finally, bicycles are concrete objects and intuitively easier to characterize with a finite list of features. Even so, we can see that it is a difficult task. But some concepts are much more difficult to characterize. How could we capture concepts such as 'cheat' or 'admirable' in terms of necessary and sufficient conditions?

The problems posed by necessary and sufficient conditions were stressed by many scholars throughout the years. As early as 1956, Wittgenstein (1956[1972]) pointed out how difficult it would be (if not impossible) to make such a checklist for the category 'game'. The shortcomings of the approach became especially apparent when Rosch (1978) and her collaborators showed that human beings consider category membership according to a scale rather than against a yes/no checklist. Some members rank as better examples than others, while others are considered members of a category only if they share a certain ***degree of resemblance*** with the ***best example***. Furthermore, this resemblance can be stronger for some members than for others. The work carried out by Rosch and her collaborators led to the formulation of the ***prototype theory***[40] and in some sense paved to way to other explanations of the way human beings construe concepts.

If we apply this to our bicycle example, a two-wheel vehicle propelled by human energy on which the user is sitting on a seat is undoubtedly a better example of the category 'bicycle' than an exercise bicycle or a bicycle on which the user is lying on the back. Human beings would probably still consider an exercise bicycle as a bicycle, but it would certainly not rank as the best example of the category. Similarly, it is doubtful that the motorbike would be regarded as a member of the category since there are too many features that differentiate motorbikes from bicycles: motorbikes have engines and are propelled by fuel or

40. Although this is outside the scope of this book, it should be mentioned that the original prototype theory was later formulated into a certain number of different versions in which the notion of 'prototype' and the 'degree of resemblance of members' are explained differently.

electricity; the driver is passive rather than active as far as propelling the motorbike is concerned.

Prototype theory and other related views have had and still have a major impact in many different areas. One of them affects the traditional way of accounting for concepts in definitions. In fact, prototype theory makes things slightly more complicated. Cruse (2011: 65) states: "In getting rid of necessary and sufficient features, prototype theory has thrown the baby out with the proverbial bath water; by providing an account of gradable centrality, it has lost the ability to set a boundary." How can we distinguish concepts from one another and how can we account for these differences in definitions? How can a definition include members of categories and still account for the fact that some are better members than others? How can a definition represent the fact that a boundary cannot be crossed? One way to consider the matter consists in assuming that definitions stated in terms of necessary and sufficient conditions correspond to the best example of the category, i.e. the ***prototype***.

The work of Rosch and others was not directly concerned with terminology. Undoubtedly, many concepts are delimited much more clearly in specialized fields, since relevant characteristics are often carefully chosen by experts. However, it was pointed out that this is not always the case. Temmerman (2000) argued that, while some concepts can be delineated in specific domains in terms of finite lists of characteristics, others are much more difficult to circumscribe. The author analyzed a corpus related to life sciences and showed quite convincingly that some concepts – while central – are not used with such precision by experts. In fact, many cognitive factors may prevent them from capturing concepts in a straightforward way.

Based on these observations, Temmerman suggests that specialized knowledge would be better approached from cognitive frameworks such as the ***prototype theory***. The author claims that some concepts are better characterized as '***units of understanding***' defined through a cognitive process during which knowledge is acquired in a given domain. This principle and others are part of an approach to terminology called ***Sociocognitive Terminology***. Temmerman (2000: 120) also suggests describing units of understanding in a way that is more compatible with how human beings construe them cognitively. Her model makes use of a template in which different features are recorded: core definition, historical information, facets showing degrees of essence; perspectives and intention, etc.

Other authors have also claimed that cognitive frameworks are better adapted to approach concepts in specialized domains. Among these are Faber (2012, 2014) and her collaborators who suggest importing insights from different cognitive linguistic frameworks to explain concepts and conceptual structures in the field of the environment. Their various proposals are incorporated into an approach called ***Frame-based Terminology***.

5.1.5 Multidimensionality

Different kinds of phenomena can affect concept delimitation. One of them has been amply discussed in terminology circles, especially in knowledge-based approaches, and is called ***multidimensionality***.

Multidimensionality covers two different intertwined phenomena. It can first be viewed as a classification challenge whereby concepts are organized differently according to the characteristic that is taken into consideration. The phenomenon was illustrated by Bowker (1993) with the concept 'wine'. It is customary to classify different kinds of wine according to a classic color distinction, namely white, red or rosé. However, the same 'wine' can lend itself to other classifications if different characteristics are selected: sugar content, origin, type of grape. These new classifications can parallel the first one based on color. Bowker notes:

> Despite the classical theory's claim that there is only one correct way of classifying a given concept, it is commonly accepted that people can "see the same thing in different ways." We use the term *multidimensionality* to describe the phenomenon of classification that occurs when more than one characteristic can be used to distinguish between things, and hence those things can be classified in more than one way. (Bowker 1996: 784)

The example of 'wine' given by Bowker is by no means an exception and many others can be found in specialized domains. Vehicles, for instance, can fall into different categories according to their size ('compact car', 'full-size car', 'light-duty truck', and so on); the way they are propelled ('electric vehicle', 'hybrid vehicle'); their use ('utility vehicle', 'luxury vehicle', 'family vehicle'); the objects they carry ('passenger vehicle', 'freight vehicle').

Multidimensionality also refers to a second phenomenon where the same concept can be conceptualized from different perspectives.[41] To take a simple example, consider the concept 'tomato'. In biology, it is viewed as a fruit and shares characteristics with other kinds of fruit. In the food industry, 'tomato' is placed on stands next to vegetables and is used in the preparation of sauces and salads. The food industry would probably classify it as a vegetable, a category that is nonexistent in biology but that the food industry opposes to fruit.

---

41. Similar phenomena were studied in other circles but were approached and characterized differently. Multidimensionality can be linked to different subtypes of subsenses (Cruse 2011). Some manifestations of subsenses are discussed in Section 5.2.5. In Frame Semantics, these differences in perspectives are explained by different backgrounds against which concepts are considered (Fillmore 1982).

Again, many other examples of this type of multidimensionality can be given. For instance, the concept 'carbon dioxide' can be first considered from the point of view of chemistry. It should thus be defined as a compound comprising carbon and oxygen that has specific properties. In the environment, although it remains a compound, its contribution to the greenhouse effect should undoubtedly be stressed. These different perspectives can also be reflected in the relations shared by 'carbon dioxide' with other concepts. In chemistry, 'carbon dioxide' will probably be linked to 'carbon' and 'oxygen' and perhaps with other carbon-based compounds. In the environment, 'carbon dioxide' is defined a specific concept of 'greenhouse gas' along with 'water vapor', 'methane', etc. (as shown in Figure 5.3).

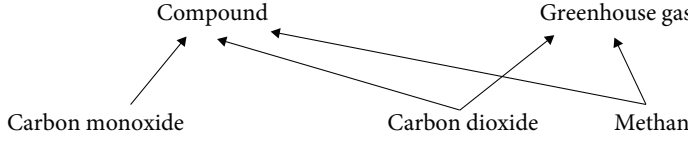


Figure 5.3 Different perspectives on 'carbon dioxide'

Since terminologists often deal with concepts from the point of view of a single subject field, the *tomato* and *carbon dioxide* cases would not necessarily represent major challenges in practice. If asked to compile a biology dictionary, terminologists would not be interested in the way tomatoes are considered in the food industry. Conversely, a terminologist working on a food industry related resource should ensure that the proper perspective is taken into consideration. The same applies to carbon dioxide: a terminologist considering this concept from the point of view of the environment should consider it as a greenhouse gas along with methane.

León Araúz and Reimerink (2010) gave a much more complex example where some concepts can display multiple dimensions in the same domain. The concept 'sand' can be defined as 'a kind of sediment located in the sea, rivers or soil layers'. However, looking at contexts in which the item *sand* appears, the authors note that the term could be associated with different series of other concepts. In geology, for instance, although still defined as a kind of sediment, 'sand' is further characterized according to grain size, and is viewed as a part of larger natural entities, such as valleys, deserts, etc. In another domain, called by the authors the *coastal domain*, 'sand' is also characterized as a part of larger natural entities, but these are restricted to coastal ones, such as beaches, and sand barriers. In addition, 'sand' is viewed as something involved in natural processes, such as waves and storms. The authors identify other differences in other domains, such as coastal defense and water treatment, noting that each of these areas trigger different conceptualizations of the concept 'sand'.

In this second situation, ***multidimensionality*** raises important issues with regard to the description of concepts in terminological resources. Should each perspective to which a concept lends itself be treated in a separate entry or term record? In other words, should all the subtle changes mentioned above about the concept 'sand' be reflected in a terminological resource? This would result in some concepts being in five, six or even more term records in which definitions and other data categories would differ only slightly. San Martín (2016) suggests an alternative solution in which multidimensionality is handled with a system of flexible definitions.

Multidimensionality raises an even more fundamental issue. How can we define a perspective formally? Do perspectives have fixed boundaries? Do they apply to sets of concepts or to individual concepts? Up until now, the literature on the subject has not provided clear answers to these questions.

Other approaches state that what matters is what the object is in essence, not the way it may be conceptualized in different situations. The presence of tomatoes in biology and in the food industry does not affect what they are in reality. Similarly, sand remains the same substance regardless of the perspective under which it is considered. Hence, according to the latter approach, terminologists should strive to define a concept according to what it is and not based on the way it is construed. This standpoint is not entirely satisfactory either. A biological characterization of 'tomato' would not allow us to capture its important features in the food industry. So the difficulty now resides in defining the "essence" of a concept with respect to a given subject field.

5.1.6 Other factors affecting the way concepts are delimited or defined

The previous section on multidimensionality explained how considering concepts from the point of view of different fields of knowledge or specific topics within the same domain can have an impact on the way they are delimited, classified and thus defined. In addition to multidimensionality, other factors can affect the way concepts are explained with definitions. Two of these factors are mentioned in this section.

The first factor is the way different kinds of users approach concepts in the same field. For instance, experts, students in environmental science, politicians, specialized translators, etc. may all find themselves in situations where environmental knowledge is required. Each group of users has a different background and hence a different understanding of central concepts in the field, such as the concepts of 'climate change', 'atmosphere', and 'circulation'. Recent studies suggest that experts and laypeople behave differently when exposed to specialized concepts (Faber et al. 2014).

Bergenholtz and Kaufmann (1997) argue that definitions should be adapted to specific kinds of users. When confronted with specialized concepts, laypeople probably need a longer definition than experts since they cannot rely on previously acquired knowledge. Similarly, students in environmental science probably need less information than laypeople, but probably more than experts.

Another important factor that affects the way concepts are delineated stems from the evolution of knowledge. Although efforts have been made in several domains to formulate consensual and standardized definitions, changes often occur over time that force experts to revise previously accepted definitions. A classic example is the concept of 'atom' that has undergone multiple revisions. Until the 1800s, it was defined as an indivisible entity. The atom was later broken down into three parts: electrons, protons, and neutrons. More recently, other elementary particles, i.e. quarks, were isolated and the structure of the atom could no longer be described in terms of three subatomic particles. Recent findings added new elementary particles, such as types of bosons and leptons. Another classic example of these changes is the recent redefinition of the concept 'planet' and the introduction of a more precise concept, i.e. 'exoplanet'.

The changes in the conceptions of 'atom' and 'planet' are now recognized and established. Scientists had to react rather quickly in order to adjust universal definitions. In many other fields of knowledge, concepts and the linguistic expressions used to label them can go through a phase of considerable instability before experts agree on a way to circumscribe them. For instance, the series of phenomena affecting climate in the world were originally labeled *global warming*. More recently, experts agreed on a more general designation, i.e. *climate change*. But 'climate change' has not yet been defined as clearly as atoms or planets, and is still being debated in different circles.

Changes in our understanding of knowledge inevitably affect the way terms are defined in specialized dictionaries. Updates must be made in order to account for more recent meanings. In some fields of knowledge, changes occur at a pace with which it can be difficult to keep up.

5.2 Lexicon-based approaches to linguistic content

Among the other perspectives under which linguistic content can be considered, we opted for a lexicon-based based approach. Let us recall some of the principles on which this approach is based and apply them to terms:

– The core entity under analysis is the ***term***; a term is defined as a ***lexical unit*** delimited syntagmatically and paradigmatically (see Section 4.3).

- A lexical unit is a lexical item with a specific ***meaning***; linguistic form and meaning are inseparable.
- The meaning of a term can be delimited according to the way it interacts with other terms and lexical units in a language.
- The meaning of a term is apprehended through a ***semasiological approach***.

Since terms are viewed as lexical units, they are an integral part of the lexicon of a language. This was pointed out by Kocourek (1991) who states that terms represent given meanings within the range of meanings a lexical item may carry. Similarly, Cabré (2003) claims that terms are units of the lexicon and that specific situational parameters "activate" a specialized value, i.e. a value that is relevant for a given subject field.

Terms are created on the basis of the lexical stock provided by language, by means of processes such as ***derivation*** (e.g. *deforestation, preprogram, reinstall*), ***neoclassical formation*** (e.g. *anthropogenic*, *eutrophication*), ***compounding*** (e.g. *keyword, groundwater, wetland*), ***composition*** (e.g. *operating system*, *greenhouse gas*); and ***addition of new meanings*** to existing lexical items (e.g. *virus* in computing; *green* in the environment). Another productive ***term creation*** method consists in borrowing items from other languages and importing them with or without adaptations (e.g. the French *bogue* is an adaptation of *bug* in English).

Since terms are an integral part of the ***lexicon*** of a language, many are recorded in ***general language dictionaries.*** Indeed, the computer terms *mouse*, *memory*, *processor* and the environment terms *greenhouse effect*, *carbon dioxide* are more than likely to be listed in recent editions of general dictionaries of English. Moreover, units that could be defined as general can be recorded in ***specialized dictionaries***: *keyboard* should normally be listed in a computing dictionary; similarly, *risk* and *vulnerable* should appear in a dictionary on the environment.

Lexicon-driven approaches are more directly concerned with understanding the place of terms within the lexicon. Section 4.4 presented four criteria that can be used to distinguish terms from other lexical units. Of course, these distinctions can only be made when separate meanings are considered. This was taken for granted rather than explained in the previous pages. The following sections show how meanings can be distinguished to help terminologists focus on the meanings that are relevant in a given field of knowledge. This being said, even if meanings can be distinguished (some more easily than others), some meanings are more closely connected than others.

5.2.1 Terms in the lexicon of a language

Since terms are part of the lexicon and adding meanings to existing lexical items is a productive method for creating terms, different meanings can coexist in a language.

In what follows, three different situations in which at least one specialized meaning is involved are described.

– **Meanings in general language/meanings in specialized domains**
A lexical item that conveys a meaning in a given specialized field can also have "general" meaning. This situation can be illustrated with the adjective *clean* that is usually associated with the meaning 'free from dirt' (*a* CLEAN *shirt*). In the environment, *clean* means 'that has a low impact on the environment' (CLEAN *fuel*). Many more examples could be given since adding a new meaning to an existing "general" lexical item is common method for creating new terms. A new meaning (often metaphorical) can be assigned to an item in the lexicon that is known to most speakers of a language as shown below with *anchor* and *taking*.

> *Anchor* (Computing): *An* ANCHOR *or link contains at least one attribute, the most common being a Hypertext Reference (HREF).*
> *(Introduction to Computers 2004)*
> *Taking* (Environment): *Taking of a species includes willfully harming an endangered or threatened animal. (Utah State University Extension Service Department of Fisheries and Wildlife 1998)*

Anchor refers to a tag placed in a web page that links it to other documents. It is based on the common meaning 'an object used to stop a boat from moving'. In the second example, *taking* covers different kinds of activities carried out by humans that can harm animals and contribute to their extinction.
The examples above show how new specialized meanings are added to existing lexical items and imported into specialized subject fields. Once created, terms can also be imported into general language. For instance, *greenhouse effect* was created to label a complex phenomenon studied by environment experts but is now commonly used in non-specialized texts. Similarly, the verb *debug*, although first used by computer scientists, is now quite common outside the realm of computer science.
The transfer of terms into general language is quite common and terms can undergo semantic changes in the process. For instance, in general language *to debug* can refer to different situations in which a problem is fixed (and not necessarily a program). Meyer and Macintosh (2000) labeled this process ***determinologization***.
The two processes described above (1. Import into a specialized field; 2. Transfer into general language) can affect the same lexical item. For instance, *bug* was created to designate an error in a computer program on the basis of the 'insect meaning' it already had in general language. Similarly, *sustainable*,

which means 'capable of being supported or upheld' in general language, has acquired a new more technical meaning in the environment, i.e. 'carried out in conformity with the protection of the environment'. Both lexical items were part of the general lexicon, acquired a specialized meaning in a given domain, and this specialized meaning migrated back into general language.

– **Meanings in different specialized domains**
A lexical item can convey meanings in two different specialized fields. Many examples were presented in previous chapters, but even more can be given. *Disinfection* in computing denotes an activity in which someone removes a virus or other malicious software from the hardware of files (*DISINFECTION of your hard disk with this program*); in the environment, it denotes an activity in which a product is used to clean or purify something (*DISINFECTION of building plumbing with chlorine*). Similarly, the French term *espace* in the environment denotes an area of a given size that is used by someone or for something (*La conversion d'ESPACES naturels en terres agricoles; En. the conversion of green areas into farming land*); in computing, *espace* can denote (among others) the portion of a storage device on which data can be placed (*l'ESPACE libre et utilisé sur un disque;* En. *available and used space on a disk*).
In some cases, a new specialized meaning can be derived from another specialized one. This is probably what happened with the term *virus.* The biological meaning served as basis to create one in computing. In other cases, two specialized meanings can be added to an existing lexical item in the general lexicon.

– **Different meanings in the same specialized domain**
Different meanings can coexist in the same subject field. For instance, the noun *download* may designate: 1. the 'process that consists in transferring a file from a remote computer to a local computer' (*a DOWNLOAD that never finishes*); 2. 'the file that can undergo or that underwent the transfer' (*a program offered as a DOWNLOAD*). Similarly, in the environment, *emission* can refer to 'the process whereby someone or something releases a substance in the atmosphere' (*EMISSION of CO_2 by men*) or to 'the substance that is sent to the atmosphere' (*toxic EMISSIONS*). The French *espace* that was discussed above has two different meanings in computing. In addition to the meaning already described, *espace* denotes a blank between two character strings (*Il faut toujours mettre un ESPACE entre le nom de la commande et les options; there must always be a space between the name of the command and the options*).

5.2.2 Criteria for semantic distinctions

When processing specialized corpora, terminologists inevitably come across lexical items that have multiple meanings. Of course, they focus more specifically

on the meanings that are relevant in the specialized domain they are asked to describe. However, in practice, ***polysemy*** needs to be considered from a broader perspective since texts contain different kinds of lexical units. Even if, in the end, terminologists fully account for meanings related to a single field of knowledge, they must often disambiguate lexical items that carry separate meanings in other situations.

Polysemy is a recurrent phenomenon in specialized texts, and terminologists need to handle it with reliable criteria. Criteria proposed in the literature on lexical semantics (e.g. Cruse 1986; Mel'čuk et al. 1995) can be used to accomplish this task. In terminology, they are used to differentiate multiple meanings that include at least a specialized one.

1. ***Substitution with a synonym*** (or near-synonym).
   According to this criterion, if a synonym can be used in all contexts in which a potentially polysemous lexical item appears, this is an indication that we are dealing with a single meaning. In contrast, if the synonym is acceptable in a subset of contexts, but not in others, then more than one meaning is associated with the lexical item. Consider the following sentences in which the adjective *green* is used.

   > *Green and pink houses in the village make perfect pictures.*
   > *My sister is wearing these awful green pants instead of the blue ones I recommended her to buy.*
   > *Eco-innovation and 'green' technologies are not just good for the environment.* *(EUROPA 2007)*
   > *… electric hybrid vehicles appear to be a green alternative to IC engines.* *(Joshi & Deshmukh 2006)*

   The near-synonym *ecological* can replace *green* in the last two sentences without altering their overall meaning, but its use in the first two is odd. This is an indication that *green* has at least two different meanings.

   > ?*Ecological and pink houses in the village make perfect pictures.*
   > ?*My sister is wearing these awful ecological pants instead of the blue ones I recommended her to buy.*
   >
   > *Eco-innovation and 'ecological' technologies are not just good for the environment.*
   > *… electric hybrid vehicles appear to be a ecological alternative to IC engines.*

   The same criterion applies to *hit* in the following sentences.

   > *To send an attachment, compose the email and then, before sending it, hit the Attachment button in your mail program.* *(Risley 2001)*

Click on the "Edit" button to review each entry, or alternatively, just HIT *the "Delete" option to remove the entry from the database.* *(Castro 2001)*

The boy promised to never HIT *his small neighbour again.*
He could HIT *people with his left hand.*

The synonym *press* can replace *hit* in the first two sentences, but not in the last two. Similarly, *punch* could replace *hit* in the last two sentences, but not in the first two. We can conclude from this that *hit* has two different meanings.

To send an attachment, compose the email and then, before sending it, PRESS *the Attachment button in your mail program.*
Click on the "Edit" button to review each entry, or alternatively, just PRESS *the "Delete" option to remove the entry from the database.*

?To send an attachment, compose the email and then, before sending it, PUNCH *the Attachment button in your mail program.*
?Click on the "Edit" button to review each entry, or alternatively, just PUNCH *the "Delete" option to remove the entry from the database.*

The boy promised to never PUNCH *his small neighbour again.*
He could PUNCH *people with his left hand.*

?The boy promised to never PRESS *his small neighbour again.*
?He could PRESS *people with his left hand.*

2. ***Differential opposition***

According to this criterion, if an antonym or another opposite is valid for all occurrences of a lexical item, then, this item conveys a single meaning. Conversely, if the opposition cannot be maintained for all occurrences of a lexical item, then, more than one meaning should be identified. This criterion supports the distinction made earlier with *green*.

Green and pink houses in the village make perfect pictures.
My sister is wearing these awful Green *pants instead of the blue ones I recommended her to buy.*

*Eco-innovation and '*GREEN*' technologies are not just good for the environment.* *(EUROPA 2007)*
… electric hybrid vehicles appear to be a GREEN *alternative to IC engines.* *(Joshi & Deshmukh 2006)*

Polluting would be a valid opposite for *green* in the last two sentences; however, it stops being a sensible choice for *green* in the first two sentences. This is an additional indication that *green* has at two different meanings in these examples. The same criterion can be used to distinguish the meanings of *extinct* in the following sentences.

Concern over species becoming rare or EXTINCT *…*
The beaver is EXTINCT *in Great Britain …*

A five week expedition into a remote EXTINCT *volcano …*
The craters of EXTINCT *volcanoes were used as reservoirs …*

Extant or *surviving* are suitable antonyms for *extinct* in the first two sentences, but not for *extinct* in the last two. Conversely, although *active* would be a possible antonym for *extinct* in the last two sentences, it would not be possible for the first two. Again, we confirm that we are dealing with two separate meanings.

3. ***Differential morphological derivation***
According to this criterion, if a morphologically related form is valid for all occurrences of a lexical item, then this lexical item conveys a single meaning. In contrast, if the morphologically related form is valid only for part of the occurrences, then the lexical item is likely to carry more than one meaning. Compare the following series of sentences with the verb *program* which was extracted from a corpus on computing.

… you have to learn how to PROGRAM *in these languages to use them.* *(Parkanski 2004)*
the tools you need to start PROGRAMMING *in Java.* *(Brain 2001)*
When you PROGRAM*, a lot of action takes place inside loops….* *(Provost 2002)*

Suppose the Network Interface Card is PROGRAMMED *to read the next packet into location 5000 in memory.* *(Brown 2001)*
The DSP is PROGRAMMED *by a set of instructions stored on another chip on the sound card.,* *(Brown 2001)*
This technology has existed for two decades, but without the PC it was hard to PROGRAM *more than a few lights.* *(Dvorak 1993)*

For the first three sentences, the following morphologically related forms can be identified: *programming* and *programmer. Programmer* would not be associated with the occurrences of *program* in the last three sentences. The last three occurrences are linked to *reprogram*, *programmable*, *reprogrammable*, items that cannot be associated with *program* in the first three sentences. These are indications that the verb *program* conveys more than one meaning.
This criterion can also apply to *biology* in the following sentences. The occurrences of *biology* in the first set of sentences can be associated with the morphologically related form *biologist* (an expert in biology). This cannot be said about the occurrences of *biology* in the last two sentences.

The reason is found in the principles of evolutionary BIOLOGY.
(Lindenmayer et al. 2011)
a multidisciplinary science organization that weaves together research on BIOLOGY, *geography, geology, geospatial information.*
(U.S. Fish & Wildlife Service 2008)

Mexican wolves were virtually eliminated before in-depth studies of their BIOLOGY *could be undertaken.*
(New Mexico Department of Game and Fish 2017)
Knowledge of the reproductive BIOLOGY *of animals is critical to understanding population dynamics.* *(Cooke 2008)*

4. ***Other differential paradigmatic relations***
This criterion verifies if other paradigmatically related items can be found for occurrences of a potentially polysemous lexical item. The criterion covers paradigmatic relations that are not covered by previous criteria. Consider the following sets of sentences in which the verb *write* is used.

To WRITE *data, the head alters the magnetic polarity of a small segment of the disk ….* *(Tyson 2001)*
Hard disks WRITE *information and read it back again straight away to make sure it's OK.* *(Brain 2001)*

Java can be used to WRITE *applications, for both Web and non-Web use.*
(Cohen 2004)
… you can use Perl to WRITE *shell scripts ….* *(Tsariounov 2004)*

The occurrences of *write* that appear in the first two sentences can be associated with a paradigm to which *read* and *access* belong. This is not the case with the occurrences of *write* in the last two sentences. They would be more naturally linked to verbs such as *program*, *develop*, and *debug*.
In Chapter 3, we used this criterion to validate that *environment* has two different meanings.

Wildlife diversity and abundance of all living things are determined by interactions among and between organisms and their physical ENVIRONMENTS.
(Utah State University Extension Service Department of Fisheries and Wildlife 1998)
Aplomados in coastal and tropical ENVIRONMENTS *are highly insectivorous ….*
(Campbell 1995)

Sustainable development thus involves protecting and improving the quality of the ENVIRONMENT. *(EUROPA 2007)*
By protecting the ENVIRONMENT, *the EU is not only tackling pressing problems, but also reflecting the wishes of its citizens.* *(EUROPA 2007)*

The occurrences of *environment* in the first two sentences can be linked to *habitat*, *range, site* and *territory,* which are all locations where species live. These relations are no longer valid for the occurrences of *environment* in the last two sentences.

5. ***Compatible or differential co-occurrence*** (Mel'čuk et al. 1995: 64–66)
This criterion is used to verify the compatibility of a term with different co-occurrents. If co-occurrents can be combined and produce an acceptable sentence, then they are compatible and the lexical item with which they are used conveys a single meaning. In contrast, when co-occurrents are combined and produce an unacceptable or odd sentence, then they are differential and the lexical item with which they are used probably carries more than one meaning.
The application of these criteria is illustrated with the examples mentioned above with *write*. *Information* and *data*, on the one hand, and *script*, *program* and *firewall*, on the other, can be combined and produce acceptable sentences. This confirms that the occurrences of *write* in the first three sentences are linked to the same meaning. However, *data* and *script* or *data* and *program*, when combined, yield abnormal sentences. This illustrates differential co-occurrence and indicates that *write* has separate meanings.

> *To* WRITE *data, the head alters the magnetic polarity of a small segment of the disk ….* *(Tyson 2001)*
> *Hard disks* WRITE *information and read it back again straight away to make sure it's OK.* *(Brain 2001)*
> WRITE *information, data on a disk*
>
> *Java can be used to write applications. for both Web and non-Web use.* *(Cohen 2004)*
> *… you can use Perl to* WRITE *shell scripts ….* *(Tsariounov 2004)*
> *Write scripts, applications*
>
> ?WRITE *information or scripts*
> ?WRITE *data or programs*

This criterion can be used to see if *environmental* has more than one meaning.

> *New PFC systems are no longer being produced due to the* ENVIRONMENTAL *impacts of these gases upon the climate ….*
> *(Intergovernmental Panel on Climate Change (IPCC) 2006f)*
> *Nonclimatic scenarios describing future socioeconomic, land use, and* ENVIRONMENTAL *changes are important for characterizing the sensitivity of systems to climate change ….*
> *(Intergovernmental Panel on Climate Change (IPCC) 2006b)*
> ENVIRONMENTAL *impacts or changes*

A greater number of people and those who are less indoctrinated seek to protect humanity, even from itself via ENVIRONMENTAL *policy.*
(Le Treut et al. 2005)

… known and respected individuals could emphasize their ENVIRONMENTAL *activities and the important role they fulfill in their lives.*
(Wołek & Wyszomirski 2013)

ENVIRONMENTAL *policy or activities*
?ENVIRONMENTAL *impact or policy*

The application of these five criteria is the perfect way to illustrate the ***relational approach*** mentioned in Section 3.1.1. The meaning of linguistic units – in this case, lexical units – is delimited according to the way they interact with others. Table 5.2 summarizes the criteria listed in this section that served to make meaning distinctions. Other terms were added to the table, but they all resort to these criteria and simply confirm a judgement on the number of meanings attached to a lexical item.

The five criteria help terminologists validate intuitions they might have about the meanings of lexical items and guide them when establishing a relationship between a lexical meaning and a specialized domain. However, even if they are extremely useful, it should be kept in mind that the perspective given by a specialized domain and the specialized corpus in which lexical items appear necessarily influence terminologists' judgements. Terminologists might make distinctions that would not be considered relevant for lexicographers since the evidence supplied by a specialized corpus might lead to more fine-grained distinctions. Different phenomena, some of which are presented further on, may slightly change the way lexical meanings are considered.

Table 5.2 Application of semantic distinction criteria

Polysemous lexical item	Criterion	Related terms
address$_1$	Differential morphological relation	*addressable, addressing*
address$_2$	Other differential paradigmatic relation	*@, username*
environment$_1$	Other differential paradigmatic relation	*habitat, range, territory*
environment$_2$	Other differential paradigmatic relation	*ecosystem*
	Differential morphological relation	*environmental*$_1$
environmental$_1$	Differential morphological relation	*environment*$_1$
	Compatible co-occurrence	*environmental policy or activities*
environmental$_2$	Substitution with a synonym	*ecological*
	Differential co-occurrence	*?environmental policy or impact*

Table 5.2 (*Continued*)

Polysemous lexical item	Criterion	Related terms
extinct$_1$	Differential antonym	*extant, surviving*
	Compatible co-occurrence	*extinct animal or plant*
	Other differential paradigmatic relation	*rare*
extinct$_2$	Differential antonym	*inactive*
	Substitution with a synonym	*active*
	Differential co-occurrence	*?extinct species and volcano*
green$_1$	Other differential paradigmatic relation	*blue, pink, white*
green$_2$	Substitution with a synonym	*ecological*
	Differential antonym	*polluting*
hit$_1$	Substitution with a synonym	*press*
	Other differential paradigmatic relation	*type*
hit$_2$	Substitution with a synonym	*punch*
	Other differential paradigmatic relation	*blow, knock, smack*
program$_1$	Differential morphological relation	*programmer, programming*
program$_2$	Differential morphological relation	*reprogram, programmable*
write$_1$	Other differential paradigmatic relation	*access, read*
	Differential morphological relation	*writable*
	Compatible co-occurrence	*write information, data*
write$_2$	Other differential paradigmatic relation	*develop, debug, program*
	Compatible co-occurrence	*write a script, a program*
	Differential co-occurrence	*?write information, a program*

5.2.3 Polysemy versus ambiguity

Polysemy overlaps with another phenomenon called ***ambiguity*** but should not be confused with it. ***Ambiguity***[42] is a property of linguistic expressions that can interpreted differently in a given context or situation. Ambiguity is especially difficult to tackle when using automated methods for processing corpora. In the sentences below (Figure 5.4), the character string *download* is ambiguous for a concordancer.

42. Another form of ambiguity affects multiword sequences. This problem, called ***structural ambiguity***, is caused by the fact that the syntactic and semantic relation between the components of a noun phrase cannot be properly identified without extralinguistic knowledge. For instance, to understand *main fuel system drain valve*, one needs to know that "*main* applies to *fuel system*, not to *fuel* or *drain valve*, and that … (the expression) refers to a valve whose function is to drain the main fuel system." (Lehrberger 2003: 213)

```
omputer, you can go to their web sites and  download  whichever one you would like to use. Anothe
hich will try to connect to the internet and  download  advertisements. I would recommend that you
ActiveX controls, your browser may try and  download  what it thinks is an animation or graphics
nternet, or on email attachments to reduce  download  time. Useful Stuff -- {STOP}-- {STOP}Most pr
ers its optional NaturalText add-on as a free  download  for owners of NaturallySpeaking Preferred o
ware, which most manufacturers offer as a  download  from their sites. Wireless Communities A nu
```

Figure 5.4 Concordances with the character string *download*

The first three concordances show occurrences of the verb *download*; the last three contain the noun *download*. Two different meanings can be associated with the noun: 1. 'the activity that consists in transferring a file from a remote location to a local one'; and 2. 'a file that has been or can be transferred from a remote location to a local one.' *Download* may be ambiguous for a concordancer, but human beings rarely find cases like these ambiguous. They can easily distinguish verbs from nouns, and the two meanings of the nouns can be identified with some of the criteria presented in Section 5.2.2.

However, some sentences might contain items that even human beings cannot disambiguate. Consider the following three occurrences of the noun *download* and compare them with the following sets of examples:

The user waits for the next DOWNLOAD.
Place the file in your DOWNLOAD *directory*.
Public DOWNLOAD *files*.
(activity or entity?)

download$_1$, n
The DOWNLOAD *order might not be chronological.*
The DOWNLOAD *speed might be 1,500 kilobits per second.*
The time needed to complete the DOWNLOAD.

download$_2$, n
In many cases, the DOWNLOAD *will be an executable file.*
File provided as a separate DOWNLOAD.

The first three examples contain attestations of the noun *download* that are ambiguous. The meaning expressed in each case can correspond either to the activity or the entity. In contrast, the last five examples are not ambiguous. Hence, polysemous lexical items may be ambiguous in a context and non-ambiguous in others. Ambiguity affects certain contextual instances of a given linguistic expression. Terminologists will avoid giving too much consideration to ambiguous cases, and will attempt to solve problems related to polysemy. They will observe several occurrences of a lexical item and make decisions based on large samples.

5.2.4 Dealing with multiple meanings

Applying the criteria listed in Section 5.2.2 allows us to distinguish different meanings attached to the same lexical item. When combined with the criteria for term identification, they can help us isolate specialized meanings and deal with them separately.

When multiple meanings are associated with one specific domain, each of them is associated with a separate description (typically, a definition) in a terminological resource. In section 5.1.2 we saw that a computing dictionary lists two definitions for *key* and a law dictionary gives four definitions for *acceptance*.

Of course, general language resources have more polysemous lexical items and more meanings to manage than specialized ones. Figure 5.5 shows the meanings listed for the noun *bank* in WordNet. Some of these meanings can be considered to be specialized as, for instance, the 'flight maneuver' that appears at the bottom of the list.[43]

Noun

- S: (n) **bank** (sloping land (especially the slope beside a body of water)) *"they pulled the canoe up on the bank"; "he sat on the bank of the river and watched the currents"*
- S: (n) depository financial institution, **bank**, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities) *"he cashed a check at the bank"; "that bank holds the mortgage on my home"*
- S: (n) **bank** (a long ridge or pile) *"a huge bank of earth"*
- S: (n) **bank** (an arrangement of similar objects in a row or in tiers) *"he operated a bank of switches"*
- S: (n) **bank** (a supply or stock held in reserve for future use (especially in emergencies))
- S: (n) **bank** (the funds held by a gambling house or the dealer in some gambling games) *"he tried to break the bank at Monte Carlo"*
- S: (n) **bank**, cant, camber (a slope in the turn of a road or track; the outside is higher than the inside in order to reduce the effects of centrifugal force)
- S: (n) savings bank, coin bank, money box, **bank** (a container (usually with a slot in the top) for keeping money at home) *"the coin bank was empty"*
- S: (n) **bank**, bank building (a building in which the business of banking transacted) *"the bank is on the corner of Nassau and Witherspoon"*
- S: (n) **bank** (a flight maneuver; aircraft tips laterally about its longitudinal axis (especially in turning)) *"the plane went into a steep bank"*

Figure 5.5 The noun *bank* in WordNet (2015)

FrameNet (presented briefly in Section 3.4.1) proceeds differently: the meaning of a specific lexical unit is said to evoke a given frame. This implies that if a lexical

43. There is no specific label in this WordNet entry for specialized usage, but most lexical resources indicate specialized meanings with usage labels.

item is polysemous, its separate meanings appear in different frames. For instance, the noun *bank* evokes the following frames (FrameNet 2018):

> Relational_natural_features: The Focal_feature is defined in relation to a Landmark_feature, either as a particular part, or as an immediately bordering entity. In addition to *bank*, the following lexical units evoke this frame: *coast.n, delta.n, estuary.n, foothill.n, mouth.n, peak.n, seaboard.n, shore.n, shoreline.n, source.n, summit.n.*
>
> Businesses: A Proprietor owns or runs a Business which provides a Product (which may be goods or services). This frame is also evoked by *boutique.n, business.n, chain.n, company.n, corporation.n, establishment.n, firm.n, mill.n, operation.n, paper.n, practice.n, shop.n, store.n.*

Although WordNet and FrameNet list and characterize different meanings, there is no clear indication as to how some meanings are connected. However, as can be observed in the *bank* entry taken from WordNet, some descriptions account for the form in which something presents itself:

- sloping land (especially the slope beside a body of water)
- a long ridge or pile
- a slope in the turn of a road or track; the outside is higher than the inside in order to reduce the effects of centrifugal force
- a flight maneuver; aircraft tips laterally about its longitudinal axis (especially in turning)

  According to WordNet (2016)

Connections can also be found between multiple meanings listed in terminological resources. For instance, all the definitions of *acceptance* in the Legal Dictionary (2015) describe situations in which someone accepts something (goods, money, a service) in a way that could be recognized in a formal manner. Slight variations appear in the way the acceptance takes place.

> **Acceptance n.**
>
> 1. receiving something from another with the intent to keep it, and showing that this was based on a previous agreement.
> 2. agreeing verbally or in writing to the terms of a contract, which is one of the requirements to show there was a contract (an offer and an acceptance of that offer). A written offer can be accepted only in writing.
> 3. receiving goods with the intention of paying for them if a sale has been agreed to.
> 4. agreement to pay a bill of exchange, which can be an "absolute acceptance" (to pay as the bill is written) or "conditional acceptance" (to pay only when some condition actually occurs such as the shipment or delivery of certain goods). "Acceptance" is most often used in the factual determination of whether a contract was entered into.

Meanings that are distributed across general language and specialized domains can also share semantic components. The ecological meaning of *green* is a metaphorical extension of the meaning of *green* as a color. Similarly, the computing meaning of *mouse* is inspired from shape of the animal. Many more examples could be given, but terminological resources – even those resources that account for multiple meanings in a given domain – do not account for the way meanings are connected in an explicit manner.

In contrast, many lexical resources devote a lot of energy to showing the different relationships shared by the meanings attached to a given lexical item. An initial distinction is made between homonymy and polysemy. ***Homonymy*** refers to a situation whereby the same lexical form conveys different meanings that cannot be connected in any way. ***Polysemy*** is also a phenomenon in which the same lexical form has multiple meanings. However, polysemous items usually share at least one semantic component.[44]

In many general language dictionaries, homonyms are recorded in separate entries while the different meanings of a polysemous item are grouped in the same entry. This is shown in Figure 5.6 with the noun *bank* taken from the Merriam-Webster's Online Dictionary (2018). Two entries are given for the noun (in addition to another one for verbal meanings). The first entry (labeled *bank I* for the purpose of this discussion) accounts for meanings that are connected to a geological formation. The second one (labeled *bank II*) describes meanings that are linked to a place where objects are held and managed. Hence *bank I* and *bank II* can be regarded as ***homonyms***. We can also see that *bank I* and *bank II* are polysemous items.[45]

In some cases, the distinction between homonyms and polysemous items can be quite difficult to make. Diachronic criteria can be of help: when no diachronic link between the meanings of an item can be established, items are said to be homonyms rather than polysemous items.

Homonymy is much less common than polysemy and the latter can take different forms. For instance, as can be seen in the entry for *bank I* from the Merriam-Webster Online Dictionary (2018), it is evident that the two following meanings are closely related: 'a piled-up mass of cloud or fog' and 'an undersea elevation

---

44. It was mentioned earlier that knowledge-based approaches have their own definitions of homonymy and polysemy. According to these approaches, homonyms are identical designations that denote different concepts that belong to separate fields of knowledge. Polysemous items are identical designations that denote different concepts in the same field.

45. We saw earlier that some lexical resources do not make a formal distinction between homonyms and polysemous items. In WordNet, all the meanings of the noun *bank* (Figure 5.5) are listed but no additional subdivision is introduced.

bank, noun

1	:	a mound, pile, or ridge raised above the surrounding level: such as
	a:	a piled-up mass of cloud or fog – a fog *bank* – a *bank* of dark clouds
	b:	an undersea elevation rising especially from the continental shelf
2	:	The rising ground bordering a lake, river, or sea or forming the edge of a cut or hollow – We lived along the banks of the Mississippi River.
3	a:	a steep slope (as of a hill) – climbed a steep bank up to the cabin
	b:	the lateral inward tilt of a surface along a curve – The engineers hadn't given the road enough *bank*; he lateral inward tilt of a vehicle (such as an airplane) when turning – The bomber crossed the target area in a sharp *bank*.
4	:	a protective or cushioning rim or piece

bank, noun

1	a:	an establishment for the custody, loan, exchange, or issue of money, for the extension of credit, and for facilitating the transmission of funds – paychecks automatically deposited into the *bank* – went to the *bank* to make a withdrawal – open a *bank* account
	b:	obsolete: the table, counter, or place of business of a money changer
2	:	person conducting a gambling house or game; specifically: dealer
3	:	a supply of something held in reserve: such as
	a:	in games: the fund of supplies (such as money, chips, or pieces) held by the banker (see [1]banker 2) or dealer
	b:	in games: a fund of pieces (such as dominoes) from which the players draw – select another domino from the bank
4	:	a place where something is held available – memory *banks*; *especially*: a depot for the collection and storage of a biological product – a blood *bank*

Figure 5.6 Entries for the noun *bank* in the Merriam-Webster Online Dictionary (2018)

rising especially from the continental shelf'. However, the following meaning 'the lateral inward tilt of a surface along a curve' is less directly connected with the first two. We can make similar observations based on the following three informal characterizations of *green*:

'of a specific color resulting from a mixture of blue and yellow'
'covered with vegetation'
'concerned with the protection of the environment'

Although all these meanings are connected, the first one is intuitively perceived as more directly linked to the second one than to the third. The connection is established on the basis that vegetation is typically green. The third one is also

construed as more closely linked to the second one than to the first, as protecting the environment is strongly associated with the idea of protecting nature.

General language dictionaries make use of complex systems in an attempt to represent the way meanings are connected as well as the distance between some of these meanings: different numbering systems; graphical devices; hierarchies, etc. The Merriam-Webster Online Dictionary (2018) uses a combination of numbers and letters which implicitly indicate that numbered meanings are more remotely connected than those introduced by letters.

A more explicit scale was introduced by Explanatory Combinatorial Lexicology (ECL) (Mel'čuk et al. 1995: 167) which accounts for three different forms that polysemy can take.

- ***Long-distance polysemy***
  This kind of polysemy occurs between a first meaning and a metaphorical extension. It can be observed between *tool*$_1$ 'An instrument used to accomplish a task' (*the hammer and the saw are* TOOLS). and *tool*$_2$ 'a method used to accomplish something' (*this is a good marketing* TOOL). *Mouse* as an animal and *mouse* as a computer peripheral would also be considered to be in a long-distance polysemy situation.
- ***Short distance polysemy***
  In this case, one or a small set of semantic components are shared by the lexical units. For instance, *ring* can designate a circular object that is used for holding something (*a towel* RING). It can also denote another kind of circular object usually made of precious metal that one wears on a finger (*a wedding* RING). These two lexical units share the components: 'object', 'circular', and 'used for something'.
- **Regular polysemy**
  The last form of ***polysemy*** called ***regular*** was defined by Apresjan as follows:

> Polysemy of the word A with the meanings a_i and a_j is called regular if, in a given language, there exists at least one other word B with the meanings b_i and b_j, which are semantically distinguished from each other in the same way as a_i and a_j and if a_i and b_i, a_j and b_j are non-synonymous.
>
> (Apresjan 1973: 16)

Regular polysemy complies with certain patterns that were identified by lexical semanticists: place versus group or people (e.g. *office* as 'a location' or *office* as a 'group of people'; contents versus container (e.g. *glass* as 'a container' or *glass* as 'the contents to be drunk'); activity versus result (e.g. assembly as 'an activity that consists in putting together separate parts' or assembly as 'a set of parts assembled'), etc. In some cases, the boundary between regular polysemy and other

kinds of meaning modulations can be difficult to establish (see Section 5.2.5 on meaning modulations). Lexical semanticists and lexicographers might disagree on the way to handle these cases. A famous example is *book* that can be regarded as a concrete object (*this* BOOK *is heavy*) or as written content (*this* BOOK *is very interesting*).

The fine-grained characterization of polysemy that was introduced above has not been used in terminology. Even when separate meanings are distinguished, they are simply listed with no real attempt to show how some of them are linked. However, a closer look at some polysemous specialized lexical items reveals that connections can easily be made.

Examples given in previous pages illustrate a very common case of ***regular polysemy*** in specialized domains: the first term conveys an activity meaning; the second one, expressed a result. The noun *download* is defined as an ***activity***: 1. the 'process that consists in transferring a file from a remote computer to a local computer'; or as a ***result***: 2. 'the file that can undergo or that underwent the transfer'. Similarly, *emission* refers to an activity: 1. 'the process whereby someone or something sends substance in the atmosphere'; or to the result of this activity: 2. 'substance that is sent to the atmosphere'.

Another case of regular polysemy found in the environment concerns a first term that denotes a whole and another one that denotes a part: e.g. sea_1 ('the part of the Earth covered with salted water'; *containers can be transported by* SEA), sea_2 ('a specific body of salted water'; *warm water in coastal* SEAS); land_1 ('the surface of the Earth and all its natural resources' (Merriam-Webster Online Dictionary 2017); LAND *and ocean absorb some emissions*); land_2 ('a delimited area of ground or soil'; *agricultural* LAND).

Another manifestation of regular polysemy affects verbs as shown in the examples below with *erode* and *warm*. *Erode* as well as *warm* denote processes in which the state of something changes (*mountains* for *erode* and *climate* for *warm*). In contrast, in the second examples in which these two verbs appear, something causes the process to occur (*wind* ERODES the soil; *forcings* WARM the atmosphere). This type of regular polysemy is a form of ***alternation*** that affect verb meaning and is sometimes referred to as an ***inchoative/causative alternation***.

As mountains ERODE*, their roots rise and are eroded in turn.* *(PANACEA 2015)*
Nowadays the wind easily ERODES *the soil.* *(PANACEA 2015)*

Western Canadian Arctic and Siberia, have WARMED *dramatically* *(Environment Canada 2007)*
Positive radiative forcings tend to WARM *the Earth's surface and lower atmosphere* *(Intergovernmental Panel on Climate Change (IPCC) 2006b)*

Another common form of alternation[46] in specialized fields is illustrated below with the verbs *print* and *simulate*. Here, we observe that, in some instances, someone carries out an activity with the help of something. This first situation is the one expressed in the first sentences in which *print* and *simulate* appear. However, in the second sentences an instrument[47] seems to be carrying out the activity *(ink jets* in the case of *print* and *models* for *simulate*). Although it can be inferred that these instruments require some sort of human intervention, linguistically the human agent is not expressed.

You can download to disk, email, or PRINT *the Web page on the IE 6 screen.* *(Cohen 2004)*
… the Ink jets now PRINT *great color photographic images ….* *(Dvorak 1993)*

Modellers undertaking these simulations first apply historical changes in greenhouse gas and aerosol concentrations to SIMULATE *how climates have changed over the past century or so.* *(Environment Canada 2007)*
General circulation models SIMULATE *the climate system using mathematical equations that describe the Earth's radiation budget, its translation into heat and motion, and the operation of the water cycle.* *(Environment Canada 2007)*

Another type of alternation is illustrated below with the verb *store*. As in the examples with *print* and *simulate*, a human agent is mentioned in one case but not in the other. In the first sentence, *store* denotes an activity in which someone places something in a given location; in the second sentence, a location contains something.

you probably use the Save As command in Word to avoid STORING *documents under C:\MY DOCUMENTS.* *(Ortiz 2001)*
This buffer STORES *data temporarily ….* *(Patz 1997)*

Finally, some verbs can display two different kinds of alternations. This phenomenon is illustrated below with *leach*. In the first sentence, a substance is released in the environment, but no explicit cause is mentioned. In the third sentence, however, the cause is expressed (*by this water*). The second sentence is related to the other two, but in this case, the substance is leached from a product that contains it. A second example is given with the verb *compile* in computing. The first sentence

46. The causative/inchoative alternation is generally recognized as a case of polysemy; however, other forms or alternations are not always considered to be manifestations of polysemy.

47. This form of alternation is sometimes called the **Agent/Instrument** alternation. Semantic roles such as **Agent** and **Instrument** will be introduced in Chapter 6.

contains an inchoative attestation of the verb; in the second case, an **Instrument** carries out the action; finally, in the last case, an **Agent** carries out the action with the help of an **Instrument**.

Bacteria present in the manure may LEACH *into the groundwater.* *(Cabral 2010)*
Plastics and metals degrade very slowly over time and can LEACH *harmful chemicals into the environment.*
(California Ocean Protection Council and California Ocean Science Trust 2011)
… salts are then LEACHED *from the root zone by this water.* *(Savariar 2014)*

… the algorithm will not COMPILE *….* *(Field 2004)*
… the GNU software and libraries COMPILE *and run the kernel.* *(LINUX Journal 2004)*
Compile the program with the Java compiler to create a Java Applet. *(Brain 2001)*

5.2.5 Meaning modulations

Besides polysemy, another much more subtle change can affect lexical items in specialized texts. A sort of "deviation" from more common meanings is perceived, but the deviation is not deemed strong enough to be regarded as polysemy *per se*. The reasons for this difference are not easy to pin down let alone characterize formally.

A first attempt to explain this phenomenon will use the verb *warm*. Some occurrences were retrieved from a specialized corpus on climate change and are compared to examples given in FrameNet (2018).[48] (It is assumed that the latter examples reflect general usage.)

Corpus on climate change:
As increases in other greenhouse gases WARM *the atmosphere and surface, the amount of water vapour also increases.* *(Environment Canada 2007)*
… releasing into the atmosphere gases such as carbon dioxide (CO_2) which WARM *the Earth's surface ….* *(EUROPA 2007)*
Energy from the sun WARMS *the Earth's surface* *(EUROPA 2007)*
This is because the oceans and atmosphere carry heat to the poles, ensuring a more comfortable balance (for humans at least), by cooling temperatures at the equator and WARMING *them at the poles.* *(UNESCO 2006)*

48. Readers are reminded that the sentences that appear in FrameNet are extracted from the British National Corpus (BNC).

A selection of examples given in FrameNet for warm (that evokes the CAUSE_TEMPERATURE_CHANGE frame) in FrameNet (2017):
Winds will remain light and variable, and this will allow all that sunshine to WARM *the air nicely, temperatures getting up to a comfortable ten or eleven celsius, that's fifty or fifty Fahrenheit.*
But each morning, when the entire team WARM *their hands on steaming mugs of tea in the kitchen at Foulrice Farm, spirits will be lifted by thoughts of a date with destiny on March 18*
The cooking had to be done on an open fire and, when we wanted a bath, we WARMED *the water in a pan and poured it into a tin bath in front of the fire.*
The anomalous heat is thought to be due to solar energy, which penetrates the surface ice in spring and WARMS *the depths.*
The glue should be spread on one surface only and if the material is cold to the touch, it should be gently WARMED *to ensure a better bond.*

We can all probably agree that *warm* conveys the same general meaning in all the sentences presented above even if some were extracted from a corpus on climate change and others from the British National Corpus. All attestations of the verb can be loosely explained as follows: 'cause the temperature of something to rise'. A closer look at both series of sentences reveal, however, that *warm* in climate change is linked to specific types of causes (*energy, increase*, *carbon dioxide, gas*). Additionally, the entities undergoing the process of warming belong to a closed set (*atmosphere, surface*). Although *warm* is used in the same way in the sentences taken from FrameNet (*solar energy that* WARMS *the depths, sunshine* WARMS *the air*), other sentences indicate a much wider range of uses. Human agents can cause the warming (*we, team*). Instruments can be used (*a pan*, for example), a source different from the cause can also be involved (*on steaming mugs*).

The phenomenon also affects the verb *introduce.* In the environment, it denotes an activity whereby someone places a species in an area where it can live and reproduce (*predators have been* INTRODUCED *at some sites*). It is linked to terms such as *reintroduce*, *introduction, colonize* and *inhabit.* It is also opposed to *eliminate* and *extirpate*. In the general lexicon, *introduce* covers of much broader range of uses and includes many activities in which someone puts something in a given place. It can be associated with lexical units such as *place*, *insert*, *insertion* and opposed to *remove*. Again, as with *warm*, a difference can be perceived between specialized usage and general ones, but it would be difficult to claim that we are dealing with a completely new meaning.

Similar phenomena are described in Cruse (2011) under the general label ***subsense***. The one that concerns us here is a subtype called ***micro-sense***. Cruse (2011:108) argues that a lexical item such as *knife*, lends itself to different readings that are mutually exclusive: $knife_1$: 'a piece of cutlery'; $knife_2$; 'a weapon'; $knife_3$:

'a surgical instrument'; *knife*$_4$: 'a do-it-yourself tool'. For each reading, a different superordinate unit and different hyponyms and co-hyponyms can be identified.

knife$_1$	superordinate	*cutlery*
	co-hyponym	*fork, spoon*
	hyponym	*bread knife*
knife$_2$	superordinate	*weapon*
	co-hyponym	*gun, cosh*
	hyponym	*flick knife*
knife$_3$	superordinate	*surgical instrument*
	co-hyponym	*scalpel, forceps*
knife$_4$	superordinate	DIY tool
	co-hyponym	hammer, chisel

(Cruse 2011: 108)

Frame Semantics provides another kind of explanation for the phenomena we are currently describing. According to the framework, some situations can be conceptualized differently in specialized domains and everyday situations. Fillmore comments on the example of 'innocent' and 'guilty' as follows:

> In both everyday language and legal language there is a contradictory opposition between INNOCENT and GUILTY. In everyday language, the difference depends on whether the individual in question did or did not commit the crime in question. In legal language, by contrast, the difference depends on whether the individual in question has or has not been declared guilty by the court as a result of a legal action within the criminal system. (Fillmore 1982: 127)

Perhaps, another way to say this would be that the background knowledge used by legal experts differs from that gathered by laypeople when dealing with what could be considered similar situations. This assumption does not contradict the "micro-sense" characterization given by Cruse. Rather, it provides a way to explain the presence of micro-senses.

The examples of *warm* and *introduce* suggest that some changes can be perceived when considering units in specialized contexts as opposed to 'general' ones. Another manifestation of meaning modulations can be seen in cases where distinctions appear relevant only when considered from the point of view of a given domain. For instance, there are two different uses for the verb *hunt*[49] in the

49. This particular case generated a heated discussion when the research group ran into it. Some were convinced that we were dealing with two different meanings; while others believed

environment. One corresponds to the activity where a meat eater chases and captures other animals for food ($hunt_1$); the second designates the activity carried out by human beings who chase animals for other reasons such as leisure, this activity having a negative impact on the conservation of species ($hunt_2$). Furthermore, human beings usually resort to instruments when hunting and, in many countries, they must comply with regulations that specify when and how the hunting can take place. $Hunt_1$ is linked to terms, such as *predation*, *predator,* and $prey_V$; while $hunt_2$ is linked to *poach*, *poaching, poacher,* $capture_V$, and $fish_V$.

The phenomena that are described in this section have other consequences for the characterization of meanings. One can be seen in the relations that a term shares with others. This was briefly mentioned in the discussions about *introduce* and *hunt*. Another example can be given with the term *password*. In the field of computing, other than the fact that the password is a character string and is used in conjunction with a login, it would be difficult to argue that we are dealing with an entire new meaning compared to *password* used in other situations. However, in computing, *password* is linked to other terms and lexical units that would not be deemed relevant from the point of view of general usage.

password (Computing):
login
username
generate a ~
type a ~
prompt someone for a ~
crack a ~

There are countless cases of meaning modulations in specialized domains. These phenomena are quite challenging when accounting for them in lexicographical and terminological resources. Should we provide separate descriptions or merge meaning modulations into a single explanation? Terminologists might consider it relevant to include the verb *warm* in a resource that deals with climate change. However, it is doubtful that lexicographers would distinguish the meaning it conveys in the field of climate change from a more general meaning. Terminologists probably make distinctions that would seem overzealous to lexicographers.

The next section examines a concrete example that illustrates the kinds of challenges that terminologists face when making meaning distinctions.

---

that the two situations were not different enough to form the basis of a real distinction. This often occurs with meaning modulations.

5.2.6 Handling complicated cases: *absorb* and *absorption*

The meanings of lexical items undergo different changes for reasons that can be explained with a mixture of phenomena covered in this chapter, namely polysemy, multidimensionality and meaning modulations. This section examines two closely related units (the verb *absorb* and the noun *absorption*) and explains how they raise challenges in practice.

Consider the following groups of contexts with the verb *absorb*. The contexts were extracted from two different corpora, a corpus on climate change and a corpus on water pollution.

Water pollution:

(1) *These hard, impervious surfaces make it easier for stormwater to pick up,* ABSORB*, and carry pollutants.* *(National Ocean Service 2007)*

… the land is then used to slow runoff and ABSORB *sediments and contaminants.* *(National Ocean Service 2007)*

Climate change:

(2) *Paradoxically, oceans reduce global warming. Indeed, they largely* ABSORB *the most important of the greenhouse effect gases: carbon dioxide or* CO_2*…* *(Environment Canada 2007)*

… causing the terrestrial biosphere to ABSORB *a significant amount of carbon.* *(Intergovernmental Panel on Climate Change (IPCC) 2006c)*

(3) *… these gases* ABSORB *much of the outgoing heat energy radiated by the earth itself ….* *(Environment Canada 2007)*
However, carbon dioxide and other gases, methane and nitrous oxide (also called dinitrogen monoxide), also ABSORB *significant amounts of thermal energy.* *(Environment Canada 2007)*

(4) *The transmitted radiation is then either* ABSORBED *or reflected at the Earth's surface.* *(Environment Canada 2007)*
Shorter wavelength sunlight passes through the atmosphere relatively unimpeded, although the ozone layer does ABSORB *a lot of higher wavelength ultraviolet energy.* *(Müller Buchdahl 2000)*

The occurrences of *absorb* in all contexts have much in common. They signal the presence of two arguments (**X** absorbs **Y**) and can be roughly explained with the following definition: 'Something takes in a substance (a liquid or energy).' Similar observations can be made by looking at the contexts containing the noun *absorption*. The noun *absorption* denotes the same process in all contexts and also has two arguments (absorption of **Y** by **X**).

Water pollution:

(1) *On-lot sewage systems, likewise do not intentionally remove these chemicals though some PPCPs are adsorbed onto the soil receiving the* ABSORPTION *field's wastewater (Penn State College of Agricultural Science 2014). Bioconcentration in fish involves the uptake of chemical by* ABSORPTION *from the water only (usually underlaboratory conditions)*
(Yarsan & Yipel 2013)

Climate change:

(2) *Carbon dioxide is constantly being removed from the air by its direct* ABSORPTION *into water and by the transfer of the carbon atom to biotic substances through photosynthesis. (Environment Canada 2007)*

However, natural emissions are offset by the natural absorption processes such as the uptake of CO2 by plant photosynthesis, as well as ABSORPTION *by the oceans. (Environment Canada 2007)*

(3) *Such energy* ABSORPTION *by the greenhouse gases heats the atmosphere. (Intergovernmental Panel on Climate Change (IPCC) 2006d)*

Nitrous oxide is 296 times more efficient at ABSORPTION *than* CO_2. *(Le Treut et al. 2005).*

(4) *… the physical characteristics of the land surface, including the vegetation cover have a strong effect on the* ABSORPTION *of solar energy and on the fluxes of heat. (Environment Canada 2007)*
… the ABSORPTION *of solar radiation by the Earth ….*
(Environment Canada 2007)

However, some differences can be perceived for both *absorb* and *absorption* when the precise nature of arguments is taken into consideration.

- (1): The process involves natural entities (organisms, soil, water) and different kinds of substances (mercury, pollutants, fertilizers).
- (2): The process involves large natural entities (oceans, forests, the biosphere, water) and gases (carbon, carbon dioxide).
- (3): Gases (aerosols, carbon dioxide, greenhouse gas, methane, nitrous oxide) absorb heat, energy or radiation.
- (4): The radiation or energy is taken in by a component of the Earth (Earth, surface, etc.).

The arguments of *absorb* and *absorption* realized in the water pollution corpus concern the introduction of harmful substances into the environment (*pollutants*, *products*, etc.). The situation is much more complicated in the climate change corpus, in which arguments and their combinations differ. A short summary is given below:

Arguments in the climate change corpus:

A natural entity absorbs a gas
A gas absorbs radiation
A natural entity absorbs radiation

These combinations highlight different aspects of the greenhouse effect. When a natural entity, such as the ocean, absorbs a greenhouse gas, it contributes to reduce the quantity of greenhouse gases that accumulates in the atmosphere. When a greenhouse gas absorbs radiation, this radiation remains trapped in the atmosphere and contributes to enhance the greenhouse effect. Finally, the surface of the Earth can absorb part of the solar radiation; the other part is sent back to the atmosphere.

Are these differences significant enough to assert that *absorb* and *absorption* have separate meanings, possibly one that has to do with the introduction of harmful substances in the environment and others that express different processes in the greenhouse effect? And if we are dealing with separate meanings, how can they be distinguished? Can distinctions be based exclusively on the fact that the realizations of arguments differ from one corpus to another while the general meaning remains unchanged? These distinctions would be difficult to sustain since the same situation can apply to many other items that appear in different corpora.

We can take a closer look at lexical relations between *absorb* and *absorption* and other terms in each corpus, in other words apply the criteria examined in Section 5.2.2. Relations that are valid for certain attestations of *absorb* or *absorption* but not for the others could indicate that were are dealing with some significant differences and potentially separate meanings. The examples listed below show terms that are morphologically and semantically related to *absorb* and *absorption* that were retrieved from the corpus on water pollution.

adsorb (X adsorbs Y: adhesion of Y at the surface of X)
… if those polymers ADSORB *higher quantities of pollutants, the consequences are most likely greater.* *(Ivar do Sul and Costa 2014)*

adsorption (adsorption of Y to X)
In addition to that, ADSORPTION *of surfactant increase the hydrophobicity of the soil, as a result, removed solubilized organic will be re-adsorbed on soil surface.* *(Paria 2008)*

desorb (X desorbs from Y: the reserve process of adsorb)
Consequently, the excess molecules adsorbed on the surface DESORBED *immediately, resulting in the observed increase in outlet surfactant concentration after one pore volume.* *(Paria 2008)*

desorption (desorption of X from Y)
Surfactants application in remediation of heavy metals has commonly been studied using soil washing/DESORPTION/extraction, adsorption onto soil, and phytoremediation. *(Shah et al. 2016)*

sorb (X sorbs to Y: take in by absorption or adsorption)
In aquatic environments, persistent organic pollutants SORB (adhere or bind) to particulate organic matter. *(Swackhamer et al. 2004)*

sorption (sorption of X to Y)
One of the most pronounced examples of the SORPTION capacity of hydrous metal oxides for trace metals is found in the manganese nodules from the oceans. *(Rehan 1991)*

We can also mention that a substance can undergo different kinds of processes once it is absorbed. These processes are expressed by verbs like *dissolve* or *degrade* or nouns such as *dissolution* or *degradation*. These related terms were hardly found in the corpus of climate change and, even when they appeared, they could not be directly linked to *absorb* or *absorption*. We could establish that the attestations of *absorb* and *absorption* in the water pollution corpus carry meanings different from the ones in the climate change corpus. We could thus create separate entries in a terminological resource.

Let us now examine the occurrences of the units retrieved from the climate change corpus. We mentioned above that the arguments expressed different kinds of entities.

A natural entity absorbs a gas
Related meaning: uptake
… carbon UPTAKE by the oceans and biosphere … *(Environment Canada 2007)*
Opposites: release
On a global scale, the volcanoes RELEASE less than 1% of human emissions of carbon dioxide and hence are a minor contributor to changes in its atmospheric concentrations. *(Environment Canada 2007)*

A gas absorbs radiation
Related meaning: trap
Greenhouse gases like water vapour, carbon dioxide, methane and nitrous oxide TRAP the infrared radiation released by the Earth's surface. *(Müller & Buchdahl 2000)*

A gas that can absorb: absorbing
ABSORBING gases – A number of naturally occurring minor gases within the atmosphere, most of which are relatively transparent to incoming sunlight, absorb most of the infrared heat energy being transmitted by the Earth towards space. *(Environment Canada 2007)*

A natural entity absorbs radiation
Opposites: reflect, radiate
Snow and ice, being white, REFLECT a lot of sunlight, instead of absorbing it.
(Müller & Buchdahl 2000)
… the outgoing heat energy RADIATED by the earth itself.
(Environment Canada 2007)

As we can see, different series of semantically related terms can be identified. This evidence could be enough to lead to different entries being created in a terminological resource. However, it is extremely difficult to confirm if we are dealing with different meanings altogether or if these are meaning modulations (micro-senses or different conceptualizations involving different entities). However, from a terminological point of view, especially in a resource that accounts for relations between terms, a distinction in this case is less difficult to handle that to try to merge these uses of *absorb* or *absorption* under a single entry.

Summary

Linguistic content is approached differently in knowledge-based and lexicon-based approaches.

In knowledge-based approaches a concept is delineated according to its position in a conceptual structure. The structure tells us how a given concept differs from others. Terminologists can refer to the conceptual system of a domain for the selection of relevant characteristics. Those characteristics become relevant when they allow us to distinguish a concept from surrounding ones. This method assumes that concepts can be properly differentiated from others and that all members of a class can be defined according to a finite list of characteristics (an approach based on necessary and sufficient conditions).

Although applicable to clearly predefined concepts, necessary and sufficient conditions raise some problems for other kinds of concepts for which a checklist of features can be difficult to establish. Concepts can also lend themselves to multiple classifications depending on the characteristics that are chosen, considered from different perspectives and change over time.

Designations can label more than one concept in separate domains or in the same domain. Many terminological resources handle this situation by recording these designations in separate term records.

Lexicon-based perspectives approach this situation from the point of view of polysemy. Different criteria can be used to make meaning distinctions: substitution with a synonym, differential opposition, differential morphological derivation, differential paradigmatic relations and compatible and differential co-occurrence.

Polysemy takes many different forms, some of which are related to what terminologists who are more grounded in knowledge-driven approaches define as multidimensionality. This phenomenon affects the way concepts are related to others in specialized domains and how they are defined.

Often, new meanings are added to existing lexical items in order to create terms. These items might already be known to speakers of a language or come from a different domain. In other cases, known lexical units are used to label specialized realities but their use is more specific or constrained. Some very subtle meaning modulations affect lexical items and can be taken into consideration in terminological descriptions.

The next chapter also concerns term meaning and focuses on specific kinds of terms, i.e. predicative and quasi-predicative terms.

Further reading

Different approaches to lexical meaning, the answers they provide, but also the questions they raise are addressed in Geeraerts (2010), Kleiber (1999) and Riemer (2015). Different views on specialized meaning are discussed in Béjoint and Thoiron (2000).

Criteria for semantic distinctions are presented in Cruse (1986) and Meľčuk et al. (1995). Types of polysemy are defined in Meľčuk et al. (1995). Regular polysemy is characterized in Apresjan (1973) and modeled in Pustejovsky (1995). A cognitive interpretation of sense boundaries is given in Croft and Cruse (2004).

In this chapter, polysemy was considered from a synchronic point of view. Some authors have looked into different processes involved in the creation of new meanings in specialized domains (Meyer and Macintosh 2000). Changes in time affecting the way terms are defined are addressed in Dury (1999). Pecman (2018) argues that language offers different ways to support the construction of knowledge.

The import of Prototype Theory in terminology in argued in Termerman (2000). Muldimensionality is introduced in Bowker (1993). León Araúz (2009) is a study on a specific kind of multidimensionality, i.e. perspectives on concepts that differ according to the subject field.

Aldestein and Cabré (2002) and L'Homme and Polguère (2008) explore the relationships between terms and the general lexicon. The place of terms in general language dictionaries was analyzed by a number of researchers, among which Alonso Campos (2008) and Josselin-Leray (2005).

An important strand of terminology studies the effect of corpora (or text genre) on terms, their meaning, and the linguistic environment in which they appear (Condamines 2005).

Finally, terminological definitions are presented in most textbooks on terminology. Sager (1990) stresses their role in terminology and lists different kinds of definitions. Vézina et al. (2009) is entirely devoted to definition writing. Sager (2000) is a collection of essays on definition. A recent special issue of *Cahiers de lexicologie* edited by Polguère and Sikora (2016) contains articles that address many different topics related to definitions (definitions in ontologies, terminological definitions, definitions for language learning, etc.).

CHAPTER 6

Predicative terms, participants and arguments

Terms, like other lexical units, can be classified into different categories depending on the linguistic property that is taken into consideration. We already saw that they can belong to different parts of speech (noun, adjective, verb or adverb) and that they denote different types of concepts, i.e. entities, activities, relations, or properties. These broad categories can be further broken down into much more precise ones. Terms with different linguistic properties obey distinct linguistic rules.

Up until now, differences in linguistic properties or behavior have not been a real concern in terminology since, as was previously mentioned in this book, traditional terminology and other knowledge-driven approaches give little attention to the linguistic properties of terms. In contrast, in lexicon-driven perspectives, these distinctions are necessarily taken into consideration as linguistic properties are an integral parts of terms.

This chapter presents a fundamental distinction between types of lexical units and terms, namely predicative units and non-predicative units. It also covers different methods for representing predicative units that have been used in terminology and other frameworks. In Chapter 3, this distinction was introduced indirectly, since both Explanatory Combinatorial Lexicology and Frame Semantics have different systems to describe predicative units. Later on an additional notion will be introduced, that of 'quasi-predicative lexical unit', proposed by Explanatory Combinatorial Lexicology.

6.1 Predicative terms and other kinds of terms

Lexical units (LUs) can denote ***entities*** (concrete objects like *computer*, *printer*; animate entities such as *bird*, *driver*; substances like *methane*, *water*, etc.) but also concepts of a different nature, such as ***activities*** (*print*, *inherit*, *degradation*), ***properties*** (*light*, *harmful*, *validity*), and ***relations*** (*offspring, parent, part*, *subset*).

The meaning of LUs that denote entities versus those that designate other types of concepts is analyzed and explained differently. Consider *bird*, an LU

denoting an entity. An informal explanation of the meaning of *bird* would read as follows:

- It is a species of animal.
- It has feathers, a beak, wings, a tail and two legs.
- It can fly.
- It lays eggs.

This characterization of the meaning of *bird* states some of the intrinsic features of the animal. In other words, we use our knowledge of the object (or our perception of what the object is) and make a list of characteristics that apply to the LU.[50] This method works for many ***lexical units*** that denote entities since they are ***non-predicative***. For instance, LUs that denote natural substances (*water*, *carbon dioxide*), natural entities (*Earth*, *ocean*) and living organisms (*bird, turtle, wildlife, tree*) can be defined by stating the physical characteristics of the corresponding objects.

Consider a different kind of LU: *donate,* a verb that denotes an activity. The verb refers to a situation in which a form of giving is carried out and results in something changing owner. A more detailed characterization of the situation is given below.

- **Someone** is responsible for carrying out the activity: i.e. **someone** gives **something** to **someone**.
- **Something** is given: i.e. **something** undergoes the process of changing owner.
- **Someone** benefits from the giving: i.e. **someone** receives the thing that is given by the person responsible for carrying out the activity.

We could also state more succinctly: ***someone*** *donates* ***something*** *to* ***someone***. In other words, explaining the meaning of *donate* requires that reference be made to the ***participants*** in the ***situation*** expressed by the lexical unit. In the previous example, participants are represented with **something** and **someone**.

What was just said about *donate* also applies to other LUs or terms, such as *print* and *degradation*. 'Print' involves someone who launches the activity, something that undergoes the activity and an instrument that is used in the process. Simply put, we can state that ***someone*** *prints* **something** *with* ***something***. Similarly, degradation is a process that affects **something** (DEGRADATION *of the soil*). Other LUs denote specific properties that apply to participants. This is the case with *light*,

50. However, as we saw earlier, terminologists normally refer to characteristics that are deemed relevant to define a concept within a special subject field. In the case of bird, the ability to fly would not be mentioned in a formal biological definition.

harmful and *validity*: **something** is *light*; **something** is *harmful* to **something** or **someone**; *validity* of **something**.

Names given to items labeled *something* and *someone* in the examples above can change from one theoretical framework to another. For instance, Explanatory Combinatorial Lexicology (ECL) uses the term *actant*, whereas that Frame Semantics prefers the label *frame element* (and more specifically *core frame element* when participants are obligatory) (see Chapter 3). These preferences reflect theoretical stances on the notion of 'participant' which will not be addressed here. We will use the term ***argument*** which is widespread in linguistic literature. However, it should be pointed out that the notion of ***argument*** presented in this chapter and later on in the book is semantic in nature. Arguments correspond to slots that must be filled in order to account for the meaning of a predicative unit. If an argument is missing, the activity or property denoted by the predicative unit is not fulfilled.

Lexical units that require ***arguments*** are called ***predicative lexical units***. LUs that denote ***activities***, ***properties*** and ***relations*** are predicative. These units have meanings with slots to be filled by ***arguments***. The notion of 'predicative unit' is more naturally associated with verbs, since arguments are often realized as syntactic groups that are connected to verbs in sentences. However, adjectives, and adverbs as well as many nouns that denote activities or properties are predicative units, as shown below:[51]

green	***something*** *is* GREEN
threat	THREAT *of* ***someone*** *on* ***something***
effect	EFFECT *of* ***something*** *on* ***someone*** *or* ***something***
buy	***someone*** BUYS ***something*** *from* ***someone***
warm	***something*** WARMS
decomposition	DECOMPOSITION *of* ***something***
migrate	***someone*** MIGRATES *from* ***point A*** *to* ***point B***

Given the focus on knowledge of many terminological approaches and its methodological consequence of accounting for nouns, terminology has not specifically dealt with predicative units. However, many terms are predicative as shown below with examples taken from the fields of computing, cycling and the environment. Terminological verbs, adjectives, and many nouns are predicative in nature. In the examples, arguments are represented with variables (**X**, **Y**, **Z**), a system used to state them without adding semantic information. We will see further on that other systems can be used.

51. Of course, other LUs, such as prepositions, and conjunctions are predicative. I focus here on parts of speech that are relevant for terminology.

Computing
dynamic **X** (*DYNAMIC site, DYNAMIC document, DYNAMIC address*)
X downloads **Y** from **Z** to **W** (*DOWNLOAD a file from a website, you can DOWNLOAD these files easily*)
processing of **X** (by **Y**) (*this PROCESSING of incoming data, picture PROCESSING by computer*)

Cycling
X rides **Y** (*RIDE a bicycle to class, how I learned to RIDE this bicycle*)
oncoming **X** (*ONCOMING traffic, ONCOMING vehicles*)
safety of **X** (*SAFETY of real users, your SAFETY*)

Environment
green **X** (*GREEN energy, GREEN transportation*)
impact of **X** on **Y** (*climatic IMPACT on such emissions, IMPACT of climate change on world heritage sites*)
X pollutes **Y** with **Z** (*we are POLLUTING the atmosphere with toxic emissions, wetlands have been POLLUTED by agriculture*)
X reforests **Y** with **Z** (*REFOREST degraded lands with native tree species, tropical countries should be paid to REFOREST net land with natural vegetation*)

6.2 Defining the argument structure of a predicative term

The ordered list of semantic arguments of a lexical unit is called the ***argument structure***. Defining the argument structure of an LU or term consists in stating its obligatory participants and the canonical order in which they appear with respect to the predicative LU. The short list below shows terms that have one or up to four arguments.

One argument

X retreats (*glaciers are RETREATING*)
extinction of **X** (*EXTINCTION of some more vulnerable species*)
sustainable **X** (*SUSTAINABLE exploitation of forests*)

Two arguments

X configures **Y** (*You CONFIGURE the router*)
X is compatible with **Y** (*chips that are not necessarily COMPATIBLE with each other*)
A visit by **X** to **Y** (*your second VISIT to a website*)

Three arguments

X reforests **Y** with **Z** (*REFOREST degraded lands with native tree species*)
X traps **Y** in **Z** (*These gases are TRAPPING more heat in the atmosphere*)

Four arguments

> **X** downloads **Y** from **Z** to **W** (*you can* DOWNLOAD *files from these websites to your laptop*)

Although this may be implicit, it is worth emphasizing the fact that a predicative structure is specific to a lexical unit, i.e. a lexical item with a given meaning. Each meaning of a polysemous lexical item may correspond to different argument structures. For instance, *connect*$_1$ ('someone establishes a link with a network in order to use it') has two arguments; whereas *connect*$_2$ ('someone attaches a component to another') has three arguments.

> **X** connects to **Y** (*Anyone can* CONNECT *to either of these networks*)
> **X** connects **Y** to **Z** (*you can* CONNECT *the power supply with the mainboard*)

In some cases, the number of arguments is the same, but their nature differs. Consider the two different meanings of *write* below:

> **X** (a hardware component) writes **Y** (data) to **Z** (a storage device) (*The memory controller* WRITES *words to memory cells*)
> **X** (a programmer) writes **Y** (a program) in **Z** (a language) (*You can use Perl to* WRITE *scripts*)

The first argument of *write*$_1$ is realized as names of devices (*controller*, *read-write head*), while the first argument of *write*$_2$ is realized by terms denoting animate entities (*programmer, developer, John*). The third arguments also differs in nature: in *write*$_1$, it can be expressed by terms denoting storage devices (*memory, disk, pen drive*); in *write*$_2$, it is instantiated as names of programming languages (*Java, C++*).

Other kinds of participants can further characterize a situation, but compared with arguments, they are optional. ***Optional participants*** are called ***circumstantials*** (or ***adjuncts***). Distinguishing arguments from circumstantials is not always an easy task. What follows are some general guidelines that may be used to define arguments (the criteria listed in this section are loosely based on those given in Mel'čuk (2004)).

A. Arguments are necessary to account for the meaning of a predicative lexical unit. If, in a given situation denoted by a LU, a participant is removed, then, the situation can no longer be expressed by the LU.
Consider the verbs *install* and *download* and some sentences in which they appear.

> *The operating system is* INSTALLED *on your hard disk.*
> *(Linux Directory 2004)*

All that generally happens when you INSTALL *a program is that the necessary files are transferred from the CD to your PC's hard disk.* (Tyson 2001)
You can even run Linux directly from a CD, without having to INSTALL *anything on your PC.* (Linux Directory 2004)
… programs that users can DOWNLOAD *from Web sites …* (Morrison 2001)
How do I DOWNLOAD *a file from the internet?* (Wired Guide 2004)
Download the file to your download directory. (Why does a PC Crash? 2001)

Install is an activity carried out by someone who puts files or programs on a storage device so they can later be opened, run or edited. Three arguments are necessary to account for the meaning of the verb: the person who carries out the activity (**argument 1**); the object undergoing the activity (**argument 2**) and the place where the installing is done (**argument 3**). The situation cannot be fully captured without these three arguments. Removing the third argument would result in a different situation no longer expressible with *install*. Adding an argument would have the same effect.

Download also denotes an activity carried out by someone who places files or programs somewhere. Therefore, we can say that it has arguments similar to those identified for *install*: the person carrying out the activity (**argument 1**); the object undergoing the activity (**argument 2**) and the place where the downloading is done (**argument *n***). However, the meaning of 'download' includes a source from where files or programs are taken. So its full argument structure differs from that of 'install' and includes four arguments: the person carrying out the activity (**argument 1**); the object undergoing the activity (**argument 2**); the location where the object is taken (**argument 3**) and the place where the object is placed (**argument 4**). All four arguments must be included to account for the meaning of *download*.

B. Arguments are expressible in texts. Often, arguments are expressed as phrases that are syntactically linked to predicative units. For instance, the arguments of *cut* in the following sentences appear in dependent syntactic groups. We assume that *cut* has four arguments (**X** cuts **Y** from **Z** into **W**)[52] when considered from the point of view of computing.

[You] can CUT *[the text] and paste it here*
You realizes the first argument as a subject.
The text instantiates the second argument as a direct object

52. We will see further on that meanings considered from the point of view of special subject fields may lead to argument structures that differ from the ones they have in the general lexicon. This is another possible effect of meaning modulations (see Section 5.2.5).

… to CUT *[the text] [off the page] and [into a clipboard]*
The text realizes the second argument as a direct object.
Off the page expresses the third argument as a complement.
Into a clipboard realizes the fourth argument as a complement.

Similarly, the two arguments of *impact* in the field of the environment (impact of **X** on **Y**) appear in different syntactic groups.

… the [economic] IMPACT *[of the hurricanes]*
Economic realizes the second argument as a modifier.
Of the hurricanes realizes the first argument as a complement.

… the IMPACT *[of such changes] [on regional surface climate]*
Of such changes instantiates the first argument as a complement.
On regional surface climate realizes the second argument as a complement.

Finally, the argument(s) of predicative adjectives may appear as syntactic heads.

However, it is not a particularly CLEAN *[fuel] …*
Fuel is the syntactic head of the predicative unit *clean*.

In some cases, the relationship between the predicative LU and its arguments is indirect, as shown in the following example:

[Changes in temperature] have a large IMPACT *[on nitrate aerosol formation]*
Here the first argument of *impact* is realized as *changes in temperature*. However, it is the subject of *have* – a support verb in this sentence – and not directly linked syntactically to *impact*.

When applying criterion B, it is important to distinguish *expressible* from *expressed*. There is often a correspondence between semantic arguments and syntactic realizations in sentences. However, this correspondence is not always straightforward for two reasons.
First, semantic arguments might not always be realized in sentences and may fill different syntactic positions, as shown below with the verb *buy.*

buy: **X** buys **Y** from **Z** for **W**
The couple BOUGHT *the house from a very rich businessman*
Three arguments are expressed.
I just BOUGHT *this wonderful book*
The first and second arguments are expressed,
This item can also be BOUGHT *separately*
Only the second argument is expressed in this sentence,

The arguments of some predicative units are seldom realized or are only realized in special cases, since they are already incorporared in the meaning of

this unit. For instance, the verb *click* has three arguments (**X** clicks on **Y** with **Z**); however, the third argument – that would normally be realized with the term *mouse* – is already included in the meaning of *click*.
Second, some groups are syntactically dependent on a predicative unit, but do not correspond to arguments, as shown below.

I BOUGHT *this book for my friend*
For my friend depends syntactically on *buy*, but does not correspond to an argument.

The company recently BOUGHT *this cutting edge printer*
Recently modifies *buy* but does not correspond to an argument.

Often, participants that express time or locations are optional. For instance, the expression of time (realized by *last Monday* in the following sentence *I* INSTALLED *Open Office on my laptop last Monday*) does not correspond to an argument of *install* since omitting it does not change the situation denoted by *install* (we said earlier that *install* has three arguments). However, time and locations can correspond to arguments for other LUs. In LUs that denote motion, such as *go*, *migrate*, and *move*, locations are obligatory and hence correspond to arguments.

When defining the argument structures of predicative terms, terminologists refer to sentences in which these terms appear since arguments are often realized in the vicinity of terms. However, corpus evidence must be handled with care for all the reasons listed above and terminologists must proceed to draw generalizations from the data.

6.3 Quasi-predicative terms

So far, predicative terms have been distinguished from non-predicative ones by focusing on the former and emphasizing the fact that they denote activities or properties while other terms denote entities. At this point, another distinction regarding terms that denote entities must be introduced since some of these terms share similarities with predicative units. Consider the following examples (adapted from Polguère 2012):

Terms that denote animate entities associated with a specific activity:
He is an EXPERT *in climatology* (an expert in **X**)
Software DEVELOPER (a developer of **X**)

Terms that denote groups of entities:
Fish POPULATION *in this area* (a population of **X**)
Area covered with herbaceous VEGETATION (vegetation composed of **X**)

Terms that denote parts or divisions of entities:
The disk's READ/WRITE HEAD (a read/write head of **X**)
Hold the left mouse BUTTON (a button of **X**)

Terms that denote entities with respect to a possessor:
The HABITAT *of these species* (a habitat of **X**)
Your PASSWORD *has expired* (a password used by **X**)

Terms that denote means or instruments:
A Web BROWSER (a browser used by **X** for **Y**)
A programming LANGUAGE (a language used by **X** for **Y**)

Explanatory Combinatorial Lexicology (ECL) defines these units as ***quasi-predicative lexical units***. According to ECL, these lexical units also have arguments and their description should include a characterization of their argument structures. Quasi-predicative LUs have the following characteristics (Polguère 2012):

A. Quasi-predicative LUs are nouns that denote ***entities*** as opposed to predicative terms that denote activities, properties or relations.
B. They have a "binding meaning". In other words, they have a meaning that requires arguments.
C. The quasi-predicative nature of LUs can be evidenced by their combinatorial properties: (1) A participant can be realized as a linguistic expression that are syntactically linked to the LU (for instance, a complement, as in *HABITAT of this species*; a possessive determiner, as in *their* HABITAT; or a modifier, as in *fish* POPULATION); (2) A participant can also appear in collocations controlled by the term (as in *this species occupies the* HABITAT).

The definition of an intermediate category between truly non-predicative units (such as *animal* or *carbon dioxide*) and predicative ones (such as *download* or *impact*) allows us to handle units that are not treated in a uniform manner by other frameworks.[53] This category also allows us to take into account the properties of other lexical units that are semantically related to them, more specifically collocates (Chapter 8 will cover this matter in more detail).

53. In other frameworks, part of the LUs defined by ECL as quasi-predicative units can be defined as predicative. For instance, LUs that denote animate entities associated with a specific activity (e.g. *developer*) are often considered to be predicative.

6.4 Argument structures in specialized versus general language

Chapter 5 mentioned that new specialized meanings can be added to existing lexical items (e.g. *taking* in the field of the environment and *anchor* in computing) and that some terms undergo meaning modulations. These phenomena inevitably have consequences for the argument structure of terms. Terminologists can consider that terms have different argument structures even if their meanings are closely related to those of general language LUs. Of course, changes in argument structures can also affect quasi-predicative terms.

Firstly, the number of arguments required to account for a meaning in a given specialized domain may differ from that required in general language or other domains. For instance, in order to account for the verb *send* in computing, four arguments become necessary compared to the general meaning of *send* that requires three arguments. The fourth argument stands for the technology.

> Send: **X** sends **Y** to **Z** (*I* SENT *a letter to my friend*)
> Send (computing): **X** sends **Y** to **Z** by **W** (*someone* SENDS *you a .doc file by email*)

Secondly, the nature of the arguments can differ. This situation is in fact much more common that the previous one. We saw in Section 5.2.5 that the meaning of the verb *warm* in the field of the environment selects arguments that are much more restricted in nature than what could be observed in general language.

> warm: **X** warms **Y**
> **X** can be realized as *sun, team, we, she, 1, breath*
> **Y** can be realized as *soil, hands, water, feet, soup, skin, air, vegetable*, etc.
>
> warm (environment): **X** warms **Y**
> **X** can be realized as *carbon dioxide, water vapour, forcing, greenhouse gas, energy, perturbation*, etc.
> **Y** can be realized as *climate, atmosphere, surface, climate system, temperature*, etc.

6.5 Representing predicative and quasi-predicative terms

Defining predicative and quasi-predicative terms in resources requires a special apparatus to account for arguments. Linguists have proposed different systems to represent ***argument structures***, two of which were used above: variables (**X**, **Y**, **Z**) and very broad semantic labels (**someone**, **something**). Other systems that are implemented in terminological descriptions are described in this section. Argument structures can also be represented graphically, as in Figure 6.1 for the verb *donate*.

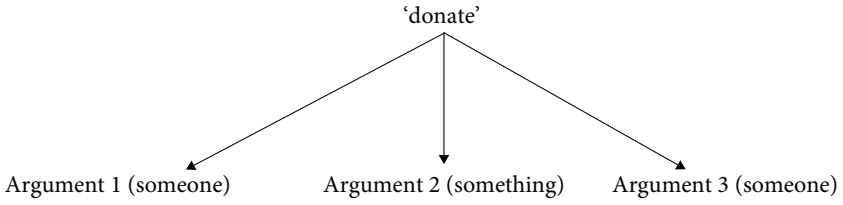


Figure 6.1 Graphical representation of the argument structure of *donate*

– ***Variables***

In Explanatory and Combinatorial Lexicology (Mel'čuk et al. 1995), arguments are represented with variables indicating that a slot must be filled by a linguistic expression.

> *X downloads Y from Z to W*[54]
> *Impact of X on Y*

The use of variables is the most neutral representation system presented in this section. It states the number of arguments and their position with regard to the predicative or quasi-predicative unit without adding semantic content.

Binon et al. (2000) used this system to label arguments in the definitions of terms recorded in a business dictionary. Although there is no separate data category to indicate the argument structure, it can be inferred from the definition that the argument structure of *achat*$_{1.1}$ (purchase) is the following: achat par **X** de **Y** auprès de **Z** (purchase by **X** of **Y** from **Z**).

> ACHAT$_{1.1}$: Opération par laquelle un agent économique (un particulier, une entreprise, une administration – **X**) reçoit un bien, une valeur un ou droit (**Y**) d'un autre agent économique (un particulier, une entreprise, une administration – **Z**) ou bénéficie d'un service (**Y**) contre paiement d'une somme d'argent.[55]

Binon et al. (2000) further explain variables with some terms that can express them. For instance, **X** is characterized as "un particulier, une entreprise, une administration" (an individual, a company, an administration).

---

54. The verb *download* (in computing) and the noun *impact* (in the environment) are used to illustrate each representation system in this section. I chose labels which might differ from what the authors cited would have chosen, but the general principles still apply.

55. En PURCHASE: 1.1 Operation in which an economic agent (an individual, a company, an administration – **X**) receives a good, a value or a right (**Y**) from another economic agent (an individual, a company or an administration – **Z**) or receives a service (**Y**) for an amount of money.

– ***Semantic roles***

Semantic roles (also called *thematic roles*, *theta roles*, or *cases*), such as **Agent**, **Patient**, **Instrument** are used to account for the deep relation between a predicative unit and its arguments, thereby capturing generalizations that surface syntax representations can miss. This was pointed out in a seminal paper by Fillmore (1968) and exemplified with the verb *break*.

> *[John*$_{\text{Agent}}$*] BROKE the window.*
> *[A hammer* $_{\text{Instrument}}$*] BROKE the window.*
> *[John*$_{\text{Agent}}$*] BROKE the window [with a hammer* $_{\text{Instrument}}$*].*
> (according to Fillmore 1968: 42)

In the first and third sentences, the subject of *break* is *John*; in the second one, the same syntactic function is held by *hammer*. In the third sentence, *hammer* is a complement of the verb. However, in all three sentences, the deep relationship between *John* and *break* and between *hammer* and *break* is the same regardless of the syntactic functions of *John* and *hammer*. Additional sentences can be added to this first set of examples in which the syntactic functions of groups expressing the arguments of *break* change: *the window was broken by the hammer; the window was broken by John, the window broke*. Again, there is no effect on the deep relationship between *break* and its arguments. In all these sentences, *window* is the argument that undergoes the breaking. *John* is the argument responsible for the activity; and a *hammer* is used to accomplish the act of breaking.

To account for the deep relation between predicative units and their arguments, Fillmore proposed a first list of roles (called *cases*): **Agentive**, **Instrumental**, **Dative**, **Factitive**, **Locative** and **Objective** that he defined as follows:

> **Agentive** (**A**), the case of the perceived instigator of the action identified by the verb, typically animate (*John* in *John opened the door* is **Agentive**).
> **Instrumental** (**I**), the case of the inanimate force or object causally involved in the action or state identified by the verb (*the key* in *The key opened the door* is **Instrumental**).
> **Dative** (**D**), the case of the animate affected by the state or action identified by the verb (*John* in *We persuaded John that he would win* is **Dative**).
> **Factitive** (**F**), the case of the object or being resulting from the action or state identified by the verb, or understood as part of the meaning of the verb (*a delicious meal* in *John cooked a delicious meal* is **Factitive**).
> **Locative** (**L**), the case which identifies the location of the spatial orientation of the state of action identified by the verb (*Chicago* in *Chicago is windy* is **Locative**).
> **Objective** (**O**), the semantically most neutral case, the case of anything representable by a noun whose role in the action or state identified by the verb

> is identified by the semantic interpretation of the verb itself; conceivably the concept should be limited to objects which are affected by the action of state identified by the verb (*the door* in *John Opened the door* is **Objective**).
>
> (Fillmore 1968: 46–48)

This preliminary set of cases (or roles) has been refined and modified over the years, but no consensual list has been established up to now. However, there seems to general agreement on the fundamental ones listed below:

- **Agent**: The argument, usually animate, that is responsible for carrying out an activity or for creating an entity (this corresponds more or less to Fillmore's **Agentive**).
- **Instrument**: The argument that refers to the object used by an agent to perform an action (this matches Fillmore's case **Instrumental**).
- **Recipient**: The argument that refers to the target of an activity carried out by an agent (this corresponds more or less to Fillmore's **Dative**, but this case can be broken down into other more specific roles).
- **Patient** or **Theme**: The argument that undergoes an activity (these two roles correspond more of less to Fillmore's **Objective**); the **Patient** is somehow affected by the activity; the **Theme** undergoes an activity but is not affected by it.
- **Location**, **Source** and **Destination**: Arguments that express different ways to consider places with respect to the predicative unit (there was only one case identified by Fillmore for locations, i.e. **Locative**); **Location** is used when no direction is involved; **Source** refers to the origin; **Destination** refers to the end point.

The roles listed above have been mainly used to account for the argument structures of verbs. They were implemented in the lexical resource VerbNet (2017) that now serves as a reference in this matter.

Semantic roles can also be used to label the arguments of terms. This is the system implemented in the DiCoEnviro (2018) and the DiCoInfo (2018). The argument structures of *download* and *impact* read as follows:

> **Agent** downloads **Patient** from **Source** to **Destination**
> **Agent:** The argument that is responsible for performing the activity.
> **Patient:** The argument undergoing the activity.
> **Source:** The argument that corresponds to the location where the patient is when the activity starts.
> **Destination:** The argument that corresponds to the location where the patient is when the activity is completed.

Impact of **Cause** on **Patient**
Cause: The argument that is responsible for the phenomenon.
Patient: The argument that is affected by the phenomenon.

– ***Object or semantic classes***

Another method for representing the arguments of predicative units is to use a system of classes that generalizes the linguistic expressions that can instantiate them. Hence, instead of listing all possible LUs that can express arguments or providing some examples (as in Binon et al. (2000)), a general label is created that is designed to include all possible realizations. This system differs from semantic roles in the sense that they do not attempt to capture the relationship between the argument and the predicative or quasi-predicative unit.

Gross (1994) developed a system of ***object classes*** (classes d'objets) that should include arguments that appear in a specific position. With object classes, the argument structures of *download* and *impact* would read as follows:

Human downloads **Representational entity** from **Remote computer** to **Local computer**
Impact of **Environmental cause** on **Environment**

Lerat (2002b) applied this system to verbs related to the field of law. The French verb *contracter* (En. take) would be described as follows:

> HUMAIN contracter (v. tr.) N <obligation> (En HUMAN takes (vt) N <obligation> (Lerat 2002b:208)

A similar method for generalizing the arguments of verbs is proposed in Corpus Pattern Analysis (Hanks & Pustejovsky 2005). A specific labeling account for types and roles. Authors claim that this is an efficient means to distinguish the meanings of polysemous verbs. For example, the difference between two of the meanings of *grasp* 'to seize hold of something' and 'to understand something' can be captured with the following descriptions:

'to seize hold of something'
[[Person=Animate]] ~ [[PhysObj]]) (one of the possible valencies)
'to understand something'
[[Person 1=Cognitive]] ~ {[[Abstract=Concept]] | [N-clause]} (one of the possible valencies)
(Hanks & Pustejovsky 2005)

– ***Frame Elements (FEs)***

Frame Semantics and the associated lexical resource FrameNet account for participants in a way that differs from the systems presented so far. Labels are used to provide information on frame elements (FEs) that are defined at the level of semantic

frames and not for lexical units *per se* (see Section 3.4.2 on Frame Semantics). In other words, they are not used to label arguments directly, but rather participants in situations that are evoked by LUs.

For example, *devour, dine, eat*, *gulp*, *ingestion*, etc. evoke the same frame (INGESTION) and thus their arguments should correspond to the following core frame elements: **Ingestor** and **Ingestibles** (FrameNet 2017).

> *[The locals* $_{\text{Ingestor}}$*] EAT [mainly fish and vegetables* $_{\text{Ingestibles}}$*].* (FrameNet 2017)
> *[He* $_{\text{Ingestor}}$*] loves to DEVOUR [huge meals* $_{\text{Ingestibles}}$*] -- washed down with Chianti wine -- which are rustled up by his wife Teresa, played by Pennies From Heaven star Gemma Craven.* (FrameNet 2018)
> *The embryonated infective eggs or larvae of the dog roundworm, [Toxocara canis* $_{\text{Ingestibles}}$*] are INGESTED [by the dog* $_{\text{Ingestor}}$*] (1) and migrate to the body tissues (2) such as the kidneys.* (FrameNet 2018)

As can be seen in these examples, the labels used are much more specific than labels for semantic roles. If we had represented the arguments of *eat* with semantic roles, we would have obtained something like **Agent** eats **Patient**. Another difference between this system and others mentioned earlier is that all participants are mentioned and labeled in a frame. They are subdivided into two categories: core (obligatory participants) and non-core (optional participants) frame elements.

In Frame Semantics, FEs are defined with respect to given frames (or within a reduced set of related frames) and this explains why most have very precise labels. According to Fillmore et al. (2003), FEs can capture generalizations about meaning that representation systems based on semantic roles, including Case grammar, would miss. For example, the argument structures of *give* and *receive* are described as follows in VerbNet (2018) which uses semantic roles:

> *Give: [John* $_{\text{Agent}}$*] GIVES [a book* $_{\text{Theme}}$*] [to Bob* $_{\text{Recipient}}$*]*
> *Receive: [Bob* $_{\text{Agent}}$*] RECEIVED [a book* $_{\text{Theme}}$*] [from John* $_{\text{Source}}$*]*

This labeling fails to capture the converse relation between *give* and *receive*. FEs, on the other hand, show explicitly the opposing perspectives profiled by the verbs: *give* focuses on the **Donor** and backgrounds the **Recipient**; conversely, *receive* focuses on the **Recipient** places the **Donor** in the background.

> *Give: [John* $_{\text{Donor}}$*] GIVES [a book* $_{\text{Theme}}$*] [to Bob* $_{\text{Recipient}}$*]*
> *Receive: [Bob* $_{\text{Recipient}}$*] RECEIVED [a book* $_{\text{Theme}}$*] [from John* $_{\text{Donor}}$*]*

This being said, the use of very specific labels for FEs does have a drawback. There are numerous labels for frame elements and their definitions must be understood within a given frame. Generalizations across large portions of the lexicon of a language might be difficult to draw.

Frame element labels have been used in terminological resources. One well-known example is a specialized resource called the *Kicktionary* (Schmidt 2009) dedicated to soccer terms. For example, in a frame called MOVE that accounts for terms such as *attack*$_V$, *attack*$_N$, *breakaway, counterattack, etc*., participants are labeled as **Team** and **Area**.

> *[PSV* $_{\text{Team}}$*] increasingly* ATTACKED *[down both flanks* $_{\text{Area}}$*], Alex and Phillip Cocu orchestrating from the back, although Panathinaikos then enjoyed a spell of possession without really threatening.*
> *It was not until five minutes from time that Gorica finally created an opportunity, but it was well smothered and inevitably led to a [Monaco* $_{\text{Team}}$*]* BREAKAWAY *with Patrice Evra forcing Mitja Pirih to save.*
> *There was almost a sucker punch in added time as [Andorra* $_{\text{Team}}$*]* COUNTERATTACKED *but Julí Sánchez shot off target.* (Kicktionary 2017)

– **Two labeling systems**

Some labeling systems presented earlier in this section can be combined to provide richer semantic information for arguments. We already saw an example with the entry *achat* from Binon et al. (2000). Variables represent arguments, but examples of terms that can realize them are also provided.

The DiCoEnviro (2018) and DiCoInfo (2018) also use two labeling systems.[56] First, arguments are labeled with ***semantic roles*** that translate the relationship they hold with the predicative or quasi-predicative term. Second, a ***typical term*** is provided, which is supposed to be representative of all the terms that can appear in a given argument position. This second labeling system was developed to make the information in the argument structure more accessible. These two labeling systems are used to describe the terms *download* and *impact* as shown below.

> Download: of **Y** from **Z** to **W** by **X**
> Semantic roles: download of **Patient** from **Source** to **Destination** by **Agent**
> Typical terms: *download of (application, file) from (computer, network) to (computer) by (user)*
>
> Impact of **X** on **Y**
> Semantic roles: impact of **Cause** on **Patient**
> Typical terms: *impact of (change) on (climate, environment)*

56. In addition to the two labeling systems, the resources give lists of linguistic realizations that are found in specialized corpora. For instance, the following list of realizations is provided for the second argument of the verb *download: antivirus program*, *applet*, *application*, *archive*, *compiler*, *copy*, *demo*, etc.

6.6 Argument structure and semantically related terms

As mentioned in the previous section, ***semantic roles*** allow us to capture generalizations that might be lost in syntax. This remains true for other labeling systems, i.e. object classes and labels for frame elements. The verb *break* used by Fillmore (1968) showed quite convincingly that even when linguistic items that instantiate arguments fill different syntactic positions, their deep relationship with the predicative unit remains unchanged.

We can take this one step forward and capture generalizations across sets of lexical units or terms that are semantically related. Some examples are provided below.

– We can capture the close relation between a verb and a noun that convey the same meaning. If the meaning is the same, it is to be expected that the number and nature of arguments are identical.

  Edit, vt: **Agent** edits **Patient** with **Instrument** (*you can* EDIT *some documents with VI*)
  Editing, n: Editing of **Patient** with **Instrument** (by **Agent**) (*The normal process for the* EDITING *of this file is using a simple editor*)

  Prey, vi: **Agent** preys on **Patient** (*these animals* PREY *on livestock, endangered species, and pets*)
  Predation, n: predation on **Patient** by **Agent** (*nest* PREDATION *by species such as pigs and fish crows*)

– Similarly, it is to be expected that synonyms and near-synonyms have similar argument structures and that their arguments are semantically related.

  Effect, n: effect of **Cause** on **Patient** (*the* EFFECT *of human activities on the composition of the atmosphere*)
  Impact, n: impact of **Cause** on **Patient** (*The* IMPACT *of volcanic forcing on climate variations*)
  Influence, n: influence of **Cause** on **Patient** (*The* INFLUENCE *of external factors on climate*)

  Program, vt: **Agent** programs **Patient** in **Material** (*the tools you need to start* PROGRAMMING *are available on the Web*)
  Write, vt: **Agent** writes **Patient** in **Material** (*you can use Perl to* WRITE *scripts*)

– Labeling arguments with semantic roles allow to show how ***alternations***[57] occur and which part of the argument structure is affected.

57. Terminologically relevant alternations were introduced in Section 5.2.4.

Inchoative – Causative
display, vi: **Patient** displays on **Destination** (*the page will* DISPLAY *on the browser*)
display, vt: **Agent** displays **Patient** on **Destination** (*When the computer needs more information it will* DISPLAY *a message on the screen*)

erode, vi: **Patient** erodes (*soft coastlines* ERODE *more rapidly*)
erode, vt: **Cause** erodes **Patient** (*the wind* ERODES *the soil*)

Agent – Instrument
print, vt: **Instrument** prints **Patient** (*this printer* PRINTS *black text quickly*)
print, vt: **Agent** prints **Patient** with **Instrument** (*You can even* PRINT *large size documents with this printer*)

– The opposition between antonyms and other kinds of opposites can also be made explicit with semantic roles. The difference between some reversive antonyms can be readily perceived as shown below with *install* and *uninstall*. One member of the pair has an argument structure with a **Source**; the other member of the pair has an argument labeled **Destination**.[58]

extinct, adj: ~ **Patient** (*it is predicted the species will be* EXTINCT *by 1995*)
extanct, adj: ~ **Patient** (*nearly one-fifth (19%) of* EXTANT *vertebrate species are threatened*)

warm, vi: **Patient** warms (*The main concern is to determine how much the Earth will* WARM *in the near future*)
cool, vi: **Patient** cools (*In a few areas, temperatures have actually* COOLED)

$install_2$, vt: **Agent** ~ **Patient** on **Destination** (*If a user decides to* INSTALL *a firewall program on a laptop …*)
$uninstall_1$, vt: **Agent** ~ **Patient** from **Source** (*you should* UNINSTALL *programs that you no longer want*)

– Finally, the use of semantic roles and typical terms also allow us to make semantic distinctions readily explicit as shown below with the verb *write* and the noun *litter*.

$write_1$: **Agent** ~ **Patient** to **Destination** (*you* WRITE *small files to a disc*)
drive, program or processor ~ data in memory or to storage device
$write_2$: **Agent** ~ **Patient** in **Material** (*you can use Perl to* WRITE *scripts*)
computer scientist ~ program in language

58. Types of antonyms are presented in Section 7.2.2.3

litter$_1$: ~ of **Agent** that contains **Patient** (*a female cheetah can give birth to a* LITTER *of one to eight cubs*).
~ of animal that contains young
litter$_2$: ~ containing **Part** in **Location** (*it covers the prey remnants with grass, leaf* LITTER*, or other debris.*)
~ of leaf in forest
litter$_3$: ~ containing **Part** in **Location** (*Plastic* LITTER *is not only ugly, it clogs streams and storm drains causing serious flooding.*)
~ of debris in environment

This method is applied in the DiCoEnviro (2018) and DiCoInfo (2018). Table 6.1 shows how it can help capture the relationships of the intransitive verb *erode$_{1a}$* with its transitive counterpart *erode$_{1b}$* and with other morphologically-related terms. The argument that undergoes the process denoted by *erode$_{1a}$* or *erode$_{1b}$* is labeled **Patient**. The **Patient** appears in the argument structure of all other semantically-related terms. Furthermore, the difference between *erode$_{1b}$* and *erode$_{1a}$*, is highlighted with the addition of a **Cause**. This is also clearly labeled in the argument structure of *erode$_{1b}$*.

Table 6.1 Argument structures of *erode$_{1a}$* and *erode$_{1b}$* and related terms

Term	**Argument structure**	
erode$_{1a}$, vi	**Patient** erodes (*the soil* ERODES)	The **Patient** undergoes the process expressed by the verb (inchoative)
erosion$_1$, n	erosion of **Patient** (*increased* EROSION *of shorelines*)	The **Patient** undergoes the process expressed by the noun. This noun has the same meaning as the intransitive *erode*
erode$_{1b}$, vt	**Cause** erodes **Patient** (*rain water is less likely to* ERODE *the limestone*)	A **Cause** is responsible for the process expressed by the verb (causative)
eroding$_1$, adj	eroding **Patient** (*ERODING coastlines are already a major threat*)	A **Patient** that is undergoing the process of erosion
erodible$_1$, adj	erodible **Patient** (*a relatively small percentage of the land is* ERODIBLE)	A **Patient** that can undergo the process of erosion
eroded$_1$, adj	eroded **Patient** (*subsurface soils on* ERODED *sites*)	A **Patient** that underwent the process of erosion

Table 6.2 shows that the principle just applied to morphologically related terms can be extended to other kinds of terms. All the terms listed in this table are linked to the verb *program* (defined as 'the activity carried out by someone that consists in writing a series of instructions'). Terms include: morphologically related

Table 6.2 Argument structure of *program* and related terms

Term	Argument structure	
$program_3$, vt	**Agent** programs **Patient** with **Material** (*the tools you need to start* PROGRAMMING *are on the Web*)	An **Agent** carries out an activity expressed by the verb; the **Patient** results from the activity; the **Material*** is used by the **Agent** to carry out this activity.
$programming_1$, n	Programming of **Patient** with **Material** by **Agent** (You can learn Java PROGRAMMING)	The noun has the same meaning as *program* shown above
$programmer_1$, n	**Agent** is a programmer of **Patient** in **Material** (*When writing the program, the* PROGRAMMER *inserts a line like this*)	This noun is the typical **Agent** performing the activity
$write_2$, vt	**Agent** writes **Patient** in **Material** (*you can use Perl to* WRITE *scripts*)	A verb with a meaning closely related to that of $program_V$
$code_4$, vt	**Agent** codes **Patient** with **Material** (*C – A programming language used to* CODE *server based applications)*	A verb with a meaning closely related to that of $program_V$
$scripting_1$, n.	Scripting of **Patient** with **Material** by **Agent** (*shell* SCRIPTING *is one of the most fun and interesting things to play around with*)	A noun with a meaning closely related to that of *programming*
$program_1$, n.	Program created by **Agent** with **Material** to act on **Patient** (*Let's say that you want to create a* PROGRAM *that prints a Fahrenheit to Celsius conversion table*)	This noun refers to the **Patient** of $program_V$
$language_1$, n.	language used by **Agent** to act on **Patient** (*programs written in* LANGUAGES *like C*)	This noun refers to the **Material** of $program_V$

*The semantic role **Material** accounts for an argument that is used by an **Agent** to accomplish an activity and that becomes an integral part of the result of this activity.

terms (*programming*, *programmer*); different parts of speech (nouns, verbs); near-synonyms (*write*, *code*); typical arguments (*language*); etc. Most of these terms require three arguments labeled **Agent**, **Patient** and **Material**.

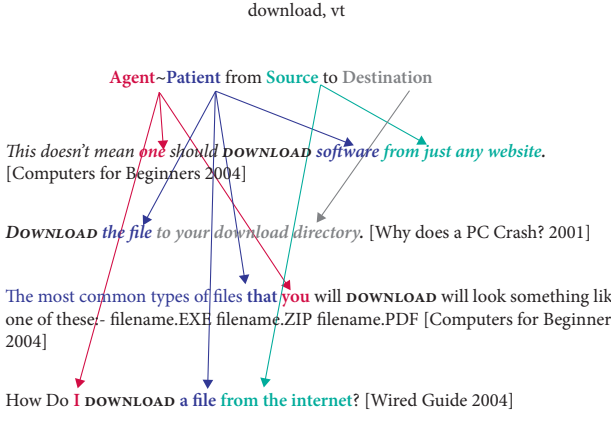
6.7 Argument structure and syntax

Section 6.2 explained that there is a connection between ***syntax*** and the ***argument structure*** that states the semantic arguments of a predicative unit. Most arguments are expressible in texts: however, their syntactic relationship with the predicative term may vary from one sentence to another. In any case, it is useful to provide descriptions of possible syntactic realizations of arguments.

As for the representation of the argument structure, different models were created to account for the connection between the semantic description of arguments and its grounding in syntax. Explanatory Combinatorial Lexicology describes the syntactic realizations of arguments in the form of government patterns (see Section 3.4.1). Frame Semantics establishes a connection between the participants in a frame (frame elements) and their realizations in sentences in the form of annotations (Section 3.4.2). Furthermore, the latter takes into consideration both arguments (core FEs) and circumstantials (non-core FEs).

The link between the argument structure and the syntactic realizations of arguments can also be represented for terms. Two terminological databases (the DiCoInfo and the DiCoEnviro) do this for part of the terms they contain, based on the methodology developed in the FrameNet project (Ruppenhofer et al. 2016). Figure 6.2 below shows how the arguments of the verb *download* are highlighted in contexts in which they appear.

The annotations highlight how the predicative term and interacts with all four arguments in different contexts. Most contexts realize only part of the arguments.



Agent	Subject(NP) (3)	one you I
Patient	Object (NP) (3) Object (Pro) (1)	software file (2) that (the most...)
Source	Complement (PP-form) (2)	website internet
Destination	Complement (PP-to)	directory

Figure 6.2 Argument structure and annotated contexts for the term *download*

The **Patient** is instantiated in all sentences. However, the **Agent** is omitted in one of them. The **Source** appears in two contexts, while the **Destination** is instantiated only once. The table next to the annotated contexts provides syntactic details about the participants: the syntactic functions (Subject, Object, etc.) held by the groups associated with a semantic role as well as the phrase type (NP, PP, etc.).

In addition to providing valuable information on the syntactic realizations of participants and the way they interact with terms, annotations help us identify semantic frames. We will come back to this in Section 8.2.2. They can also be used to spot structural divergences between equivalents in different languages.

Summary

Terms can be non-predicative, predicative or quasi-predicative. The last two categories of terms require arguments (obligatory participants) to fully capture their meaning. Two criteria can be used to describe the argument structure of a predicative or quasi-predicative term; (A) arguments are necessary to account for the meaning of the unit; (B) arguments are expressible in texts.

Arguments must be distinguished from circumstantials. The former are a core part of the meaning of a predicative or quasi-predicative unit. The latter, although they can be expressed as phrases syntactically linked to predicative units, are optional.

Different systems are used to represent the argument structure of terms (variables, semantic roles, and semantic classes). Frame elements fill a similar function although they are not designed to represent arguments *per se*. Rather, they represent participants in a conceptual situation. Semantic roles, semantic classes and frame elements capture generalizations that can escape the syntactic representations of a given lexical unit or term. Furthermore, they can capture generalizations between sets of semantically related terms.

Different methods are also used to represent the link between the argument structure and the syntactic realizations of arguments. These methods show how arguments interact with predicative units in sentences.

In this chapter and the previous one on meaning and polysemy, we saw how the meaning of different kinds of terms can be delimited, characterized and represented. The next chapter focuses on relations between terms than can be understood once their meanings are properly delimited.

Further reading

There is sizeable literature on arguments in linguistics, especially since different theoretical stances can be taken on the notion of 'argument'. The most compatible perspective with the point of view taken in this book is presented in Mel'čuk (2004).

Fillmore (1968) is a classic for anyone interested in case grammar or semantic roles. Information about VerbNet (and especially definitions of the labels for semantics roles) can be found in *VerbNet, A class based lexicon* (2017). The resource PropBank (2018) provides a similar although slightly more neutral labeling of arguments. The notion of 'object classes' are presented in Gross (1994) and that of 'quasi-predicative unit' in Polguère (2012). Corpus Patterns Analysis is discussed in Hanks and Pustejovsky (2005).

There is very little literature on predicative units in terminology. For a seminal contribution on the subject, refer to Lerat (2002a). L'Homme (2012) explains how argument structures of terms can be defined and their realizations made explicit in sentence annotations. L'Homme (2015b) argues that many terms are predicative and proposes methods to account for them.

CHAPTER 7

Relations between concepts and terms

Understanding and representing relations between concepts or terms are fundamental aspects of terminological analysis. This was mentioned indirectly in previous chapters. In Section 2.1, it was stated that the General Theory of Terminology (GTT) assumes that knowledge has a structure. Section 2.2 showed that thesauri and ontologies formally represent chosen relations between concepts. Although they do not appear explicitly in term banks, relations are taken into account by terminologists when they compile term records. In Section 3.1, the example of *program* showed that the term or the concept it denotes shares different kinds of relations with other concepts or terms. Finally, Chapter 6 contained examples of how predicative and quasi-predicative terms share argument structures.

In terminology, it is assumed that an underlying ***structure*** connects ***concepts***, in approaches that focus on knowledge, or ***terms*** and term ***meaning***, in terminological perspectives based on the lexicon. These structures result from the delimitation of sets of concepts or from the establishment of boundaries between the meanings of terms. This chapter presents a brief characterization of different kinds of relations. Chapter 8 explains how larger structures can be derived from these relations.

Knowledge-driven and lexicon-driven approaches and their respective perspectives on meaning were discussed in Chapters 3 and 5, but this chapter and the following one perhaps best highlight their differences. Knowledge-driven approaches are designed to show how knowledge items are organized in a given domain according to ***conceptual relations***. Different kinds of conceptual relations are presented in Section 7.1. Lexicon-driven approaches aim to understand relations between terms and the meaning they convey. ***Terminological relations*** are the focus of Section 7.2. The approaches are separated for the sole purpose of clarifying the perspectives taken on relations: how to define and represent them, which relations are taken into consideration.[59]

---

59. Specific terminology is also used throughout this chapter and Chapter 8 to label the properties and components of the relations defined as conceptual or terminological. It should be kept in mind that this is a personal choice and not a true reflection of the terminology used in the literature on the subject.

7.1 Conceptual relations and conceptual structures: A matter of classification

In knowledge-based approaches a concept in a special subject field is delineated according to the position it holds in a ***conceptual structure***. This was illustrated in Chapter 2 with the classification of species (Figure 2.1) and the miniature structure showing how 'mouse' is connected to other concepts (Figure 2.3).

Conceptual structures are obtained by means of ***classification.*** Organizing knowledge allows human beings to have a better understanding of the numerous and apparently unrelated realities surrounding them. Classification is achieved on the basis of shared ***characteristics***. Some concepts share a large number of characteristics while others share fewer. For instance, a 'leopard' and a 'lion' have much in common; however, a 'leopard' has much less in common with a 'cow'; 'leopard' and 'bee' share even fewer characteristics. Furthermore, concepts differ according to a varying number of characteristics: the number of characteristics that separate a 'leopard' from a 'lion' are fewer than those that distinguish 'leopard' from 'bee'. The list below shows how characteristics apply to the concepts that were just mentioned. As we add characteristics, the number of corresponding concepts diminishes.

'animal': bee, cow, leopard, lion
'animal' + 'mammal': cow, leopard, lion
'animal' + 'mammal' + 'feline': leopard, lion

We can take it for granted that characteristics that fully capture a concept are part of the reality surrounding us. In practice, however, experts or terminologists consider that some characteristics are more relevant than others. In the standard classification of animal species, characteristics chosen are based on reproduction and morphology and not other aspects of animal life, such as diet, habitat or mobility. Of course, it would be possible to propose another classification of species – one based on their habitat, for instance. This would result in a classification in which species would fall into categories such as 'marine species' and 'terrestrial species' instead of others such as mammals, insects, or birds. Even if we obtain a different distribution of animal species, we still made a selection among possible characteristics and organized concepts accordingly.

7.1.1 The backbone of a conceptual structure: The taxonomy

The most important relations in knowledge-driven approaches to terminology are taxonomical. These are believed to constitute the backbone of knowledge structures and can be found in nearly all conceptual systems. Taxonomies naturally emerge from classification.

Taxonomic relations are held by concepts that share characteristics. The two series of examples below list concepts with some of the characteristics they have in common.

> 'bicycle'; 'folding bicycle'; 'monocycle'; 'recumbent bicycle'
> Characteristics: 'small vehicle'; 'mainly man-powered'; 'has a mechanism to allow the rider to propel it'; etc.

> 'greenhouse gas'; 'carbon dioxide', 'methane', 'water vapour'
> Characteristics: 'substance'; 'in gaseous form'; 'found in the atmosphere in high concentration'; 'trap heat'; etc.

In a taxonomy, some concepts lend all their characteristics to others. For instance, all the characteristics of 'greenhouse gas' are passed on to 'carbon dioxide', to 'methane', and to 'water vapour'. 'Greenhouse gas' is then defined as the ***generic concept***; 'carbon dioxide', 'methane' and 'water vapour' are ***specific concepts*** with respect to 'greenhouse gas'. Specific concepts, on the other hand, differ from the generic one by one or a few additional characteristics. This is why taxonomical relations are defined as ***hierarchical*** and are usually represented graphically in the form of a tree. Many examples of these relations were given in previous chapters. Figure 7.1 presents a simplified taxonomy of musical instruments.[60]

In Figure 7.1, 'musical instrument' is the generic of 'woodwind instrument', 'brass instrument', 'string instrument' and 'percussion instrument'. These, in turn, are the specific concepts of 'musical instrument'. 'Woodwind instrument' can also be defined as a generic concept with respect to 'clarinet'.

It should be noted that the relations between a generic and specific concepts are ***asymmetric***. The label chosen to represent them must take into account the direction in which the relation is considered. Hence, <is-a> and <type-of> can only be used when the relation is explained from the specific concept to the generic one (a 'cello' <is-a> 'musical instrument', but not *a 'musical instrument' <is-a> 'cello').

As can be seen in Figure 7.1, taxonomies may have more than one level. In this taxonomy, three levels were identified. The addition of levels to a taxonomy does not affect the nature of the relation: in other words, the relation that holds

---

60. Since the relation shared by all these concepts is the always the same, the arrows are not labeled. However, an example was given in Figure 2.7 in which arrows are labeled <is-a>. Other labels can also be used: for instance, <type-of> (see Figure 8.5). We saw in Section 2.2 (Figure 2.6) that a thesaurus, the GEMET, establishes hierarchical relationships between concepts. However, they are labeled differently: *broader terms* are more or less generic concepts; *narrower terms* are specific concepts.

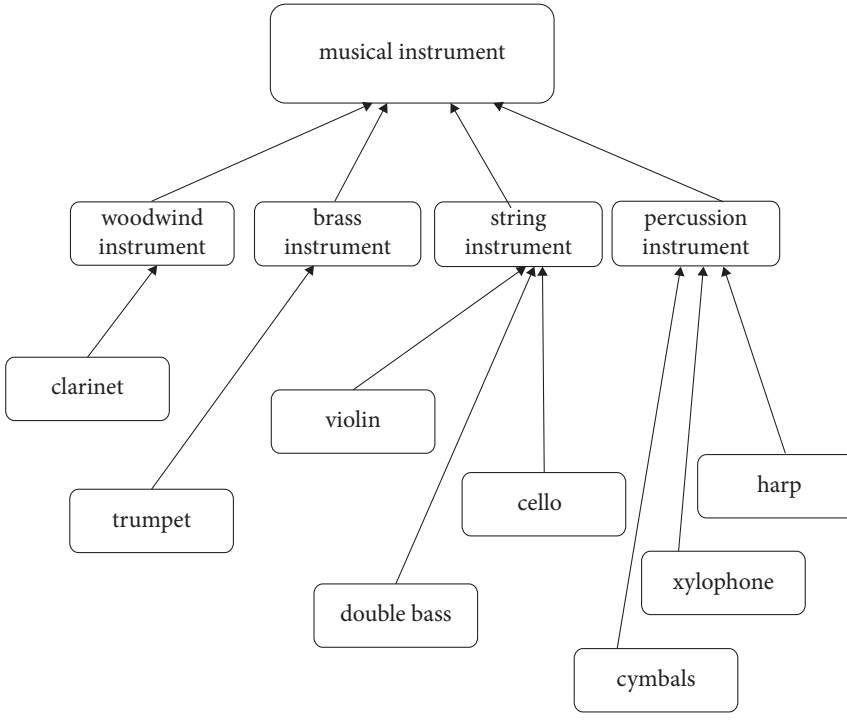


Figure 7.1 A simple taxonomy of musical instruments

between 'percussion instrument' and 'musical instrument' is the same as the one that links 'harp' to 'musical instrument'. We can say that 'a harp is a percussion instrument'; we can also say that 'a harp is an instrument'. We can go on like this if we add specific concepts under 'harp', namely 'open harp', 'arch harp', and 'bow harp'. In other words:

> 'open harp' <is-a> → 'harp' <is-a> → 'percussion instrument' <is-a> → 'musical instrument'
> 'open harp' <is-a> → 'percussion instrument' <is-a> → 'musical instrument'
> 'open harp' <is-a> → 'musical instrument'

No matter how many levels a taxonomy has, there will be no effect on the essence of the relation. This fundamental property of taxonomies is called ***transitivity***.

7.1.2 Partitive relations

The second important set of relations considered in knowledge-driven approaches are those that link parts and wholes. ***Partitive relations*** are established between

concepts that are connected spatially, or find themselves in the same vicinity. The basis on which "vicinity" is defined can vary as we will see further in this section.

The two important components of a partitive relation are: 1. a ***whole***; and 2. one or several ***parts***. In a partitive relation, concepts might or might not share characteristics, but this is not a precondition to define the relation itself as opposed to taxonomical relations where shared characteristics form the basis of the relation. The examples below show how two different concepts can be divided into parts:

Whole: 'bicycle'
Parts: 'handle bar'; 'wheel'; 'chain'; 'seat'

Whole: 'Hammer'
Parts: 'face'; 'handle'; 'neck'; 'cheek'; 'claw'

Partitive relations, as taxonomic relations, are ***hierarchical***. Graphical representations are also similar and often take the form of trees. Figure 7.2 shows the links between 'bicycle' and some of its parts.[61] 'Bicycle' represents the whole with respect to 'seat', 'handle bar', wheel' and 'frame'. These four concepts stand for the parts of 'bicycle'. Similarly, 'tire', 'spoke' and 'rim' are parts or 'wheel', which then becomes the whole with regard to these three concepts.

A representation with partitive relations can have several levels. This can be seen in Figure 7.2 where three levels were identified ('spoke' → 'wheel' → 'bicycle'). However, the transitivity property that characterizes taxonomies can no longer be observed here or can only be observed in some instances. For example, although technically, the 'spoke' can be defined as a part of the 'bicycle', it would be odd to express it that way. The 'spoke' is a part of the 'wheel', but would we say that the 'spoke' is a part of the 'bicycle'?

Partitive relations are much more complex than taxonomic ones and can be further subdivided into different subtypes. In the examples presented above with 'bicycle' and 'hammer', the parts differ from the whole (a seat differs from the bicycle) and differ from each other (a seat differs from a wheel). The parts can also be removed from the whole. This will inevitably affect the functioning of the whole, of course, but not its existence. A bicycle remains a bicycle even if a wheel was removed. This partitive relation connects a whole and a functional part.

---

61. Again, since the relation is always the same, arrows were not labeled. In some graphical representations, the relation is labeled <***part-of***>. It should be noted that it can only be expressed that way when the relation is examined from the concept standing for the part to one representing the whole. Partitive relations, as taxonomic relations, are ***asymmetric***.

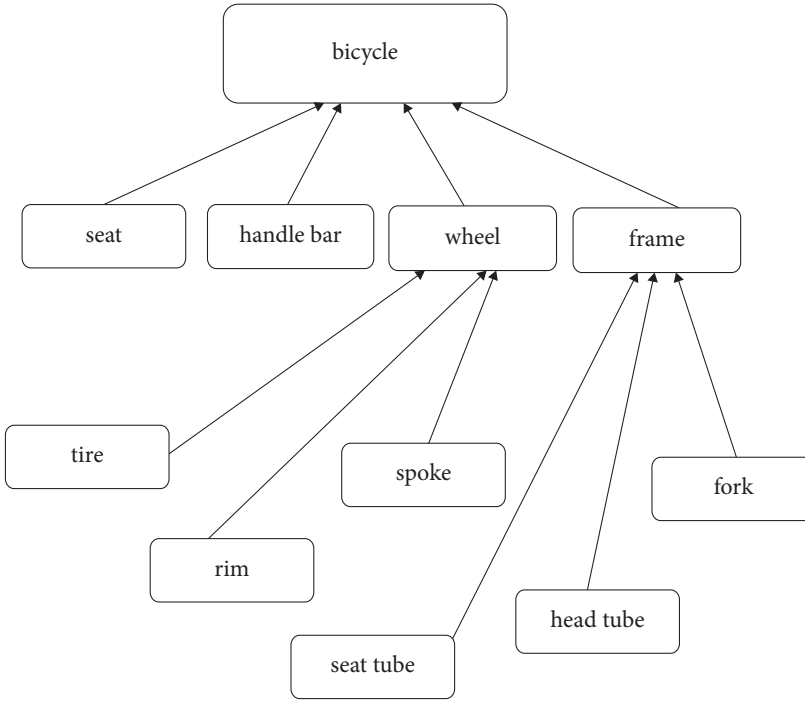


Figure 7.2 Partitive relations shared by 'bicycle' and different parts

Other partitive relations do not have the properties just mentioned. Consider the following examples:

> 'elephant' ↔ 'herd' (<member-group>): In this case, the parts differ from the whole, but all the parts are similar; a part can be removed from the whole and this will not affect the existence nor the functioning of the whole. Although different parts are perceived (many different elephants of varying shapes and ages), they are all expressed with the same designation.
> 'slice' ↔ 'bread' (<portion-whole>): Here, the part keeps all the properties of the whole even when it is physically detached from it (the slice is still bread).
> 'carbon dioxide' ↔ 'oxygen' (<constituent-whole>): The parts differ from the whole; however, in contrast with other subtypes of partitive relations, removing a part affects the existence and integrity of the whole.

Furthermore, activities divided into different stages are sometimes presented as partitive relations. For instance, 'waste management' can be broken down into 'collecting', 'sorting', 'recycling' or 'eliminating', and so on (Figure 7.3). When activities are involved, the steps are usually carried out in a specific order. Hence, the

resulting representation differs from the hierarchy reproduced in Figure 7.2 where all the parts connected to the same whole are on the same level.

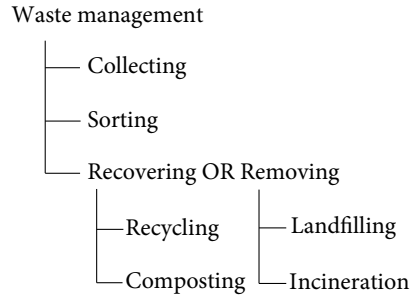


Figure 7.3 Different stages in waste management

Some locative relations can sometimes be confused with true partitive relations. For instance, there is an obvious relation between 'pen drive' and 'USB drive' since the former needs to be inserted into the latter for a user to be able to access its content. However, technically, the 'pen drive' is not a part of the 'USB drive'. Similarly, the brain is found in the cranium, but cannot be defined as a part according to the properties mentioned above.

7.1.3 Conceptual synonymy

Sections 7.1.1 and 7.1.2 examined two sets of relations shared by concepts. ***Synonymy***, also a fundamental relation, is considered from the point of view of knowledge-driven approaches as a relation between two or more ***designations***. When different designations refer to the same concept, we obtain ***exact synonymy***.

In terminological resources, exact synonyms appear on the same term record and a single definition applies to all of them. In TERMIUM Plus® (2017), for instance, the concept defined as 'A computer, usually in a computer center, with extensive capabilities and resources to which other computers may be connected so that they can share facilities' is associated with the following English designations: *mainframe, mainframe computer, central computer, main frame, main computer, main site computer*. In a conceptual structure, the relationships between designations is represented as in Figure 7.4. Chapter 2 explained that the General Theory of Terminology advocates the choice one designation for a concept. However, a milder version of this principle is applied in most terminological resources. A designation is preferred among those that are possible. In the TERMIUM Plus® example, *mainframe* was chosen (in this case, the designation is also standardized).

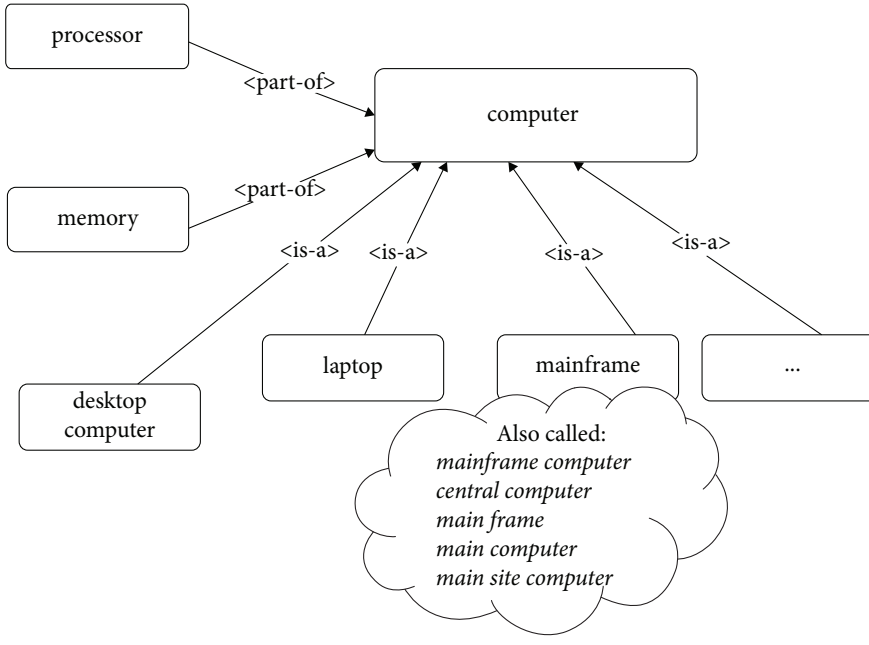


Figure 7.4 Exact synonyms in a conceptual structure

Knowledge-driven approaches, although chiefly interested in exact synonymy, can consider that some designations do not have exactly the same status as far as ***usage*** is concerned. For instance, in French, *dioxyde de carbone* (*carbon dioxide*) and *gaz carbonique* are both used to refer to the same carbon-based compound. The latter is acceptable in informal situations, whereas *dioxyde de carbone* is the adequate scientific designation.

Difference of usage is caused by various factors: different geographical areas (e.g. Canada versus Britain), a level of specialization (e.g. expert versus layperson), a time frame (e.g. obsolete versus current), or a theoretical stance (e.g. different linguistic theories). Usage can also vary according to the communication channel (oral versus written communication). Different fields of knowledge can also use specific labels to talk about the same concepts. It should be emphasized that, even if knowledge-driven approaches recognize this kind of variation, the designations are considered to have the same ability to label a concept. This phenomenon is sometimes called ***near-synonymy***.[62]

62. We will see further on that lexicon-driven approaches define near-synonymy quite differently.

In Chapter 3, another notion closely related to synonymy was introduced, i.e. ***terminological variation***. The latter notion only partially overlaps with synonymy however. It includes a wider range of phenomena where the same concepts can be expressed differently in text. ***Synonyms*** belong to the same part of speech and share structural properties. In contrast, ***term variants*** may take many different linguistic forms as shown in Table 7.1. Variants include: inflected forms; graphical variants, synonymic and near-synonymic expressions. Finally, different contextual phenomena affecting the structure of terms, such as insertion or omission, are regarded as variants as well.

Terminological variation raises challenges when attempting to identify terms in corpora especially when automated or semi-automated methods are used. Some term variants, called ***denominative variants*** (Freixa 2006), are considered to be synonyms and listed as such in a term record. Other variants can be found in running text, and might help terminologists understand a concept, but they are merely contextual and should not be recorded in term banks.

Table 7.1 Synonyms and variants for the concept 'cardiopathy'

Synonyms	Variants
cardiopathy	cardiopathy, cardiopathies
cardiac disease	cardiac disease, cardiac diseases, these diseases, they, cardiovascular disease
heart disease	heart disease, heart diseases, coronary heart disease, these diseases, this disease, heart failure, heart valve diseases
	…

7.1.4 Opposition as a conceptual relation

Since concepts are clearly delineated in conceptual structures, ***opposition*** is implicitly taken into account in knowledge-driven approaches. Concepts must be differentiated and this distinction may depend on opposing characteristics. Paradoxically, relations of opposition have seldom been addressed directly in terminology work.

Strictly speaking, a conceptual structure – and more specifically a taxonomy – accounts for ***incompatible concepts.*** Specific ***concepts*** that fall under the same generic are incompatible in the sense that they are ***mutually exclusive***. The concepts presented in Figure 7.1 under the generic 'string instrument,' i.e. 'cello', 'violin', and 'double bass' are incompatible. This means that if something is a 'cello', it cannot be a 'violin'.

Incompatibility cannot be assimilated to ***opposition*** *per se* as two concepts can be incompatible without necessarily being opposites. However, some classification systems make use of opposing characteristics, as shown below with 'manual tool' and 'power tool' and between 'output device' and 'input device'.

'tool'
- 'manual tool'
  - 'hammer'
  - 'screwdriver'
  - 'saw'
  - …
- 'power tool'
  - 'drill'
  - 'circular saw'
  - …

'peripheral'
- 'output device'
  - 'monitor'
  - 'printer'
  - …
- 'input device'
  - 'keyboard'
  - 'mouse'
  - …

7.1.5 Other conceptual relations

Taxonomic and partitive relations, even if they are still presented as the fundamental relations in terminology, do not exhaust all possible relations between concepts. In fact, in some fields of knowledge, other relations are considered even more important. Chapter 2 (Section 2.2) showed that thesauri define ***associative relations***, which encompass many different types of relations (other than taxonomic ones). More recent literature in terminology has started to further characterize other extremely important relations, such as the one that links causes and effects. In scientific disciplines, ***<cause-effect> relations*** are crucial to access knowledge.

Sager (1990) listed different kinds of conceptual relations that may be observed in various domains. The author stresses the fact that these relations might be more

important than taxonomic or partitive ones to understand certain kinds of concepts. This list is reproduced in Table 7.2, but some of the original examples were changed. In contrast with partitive and taxonomic relations, these ***relations*** are ***non-hierarchical*** (they are also ***asymmetric***) and often involve only two concepts.[63] We will see in Chapter 8 that attempting to represent many different conceptual relations, although theoretically desirable, can raise some challenges in practice.

Table 7.2 Various conceptual relations (Sager 1990: 35)

Related concepts	Concept 1	Concept 2
<cause-effect>	'greenhouse gas'	'greenhouse effect'
<material-product>*	'aluminum'	'frame'
<material-property>	'semiconductor'	'conductivity'
<material-state>	'iron'	'corrosion'
<process-product>	'manufacturing'	'product'
<process-instrument>	'editing'	'editor'
<process-method>	'recovering'	'composting'
<process-patient>	'eliminating'	'waste'
<phenomenon-measurement>	'light'	'Watt'
<object-counteragent>	'poison'	'antidote'
<object-container>	'tool'	'tool box'
<object-material>	'table'	'wood'
<object-quality>	'gas'	'high octane'
<object-operation>	'truck'	'transportation'
<object-characteristic>	'species'	'vulnerability'
<object-form>	'book'	'paperback'
<activity-place>	'farming'	'land'

*This as well as the <object-material> relation is a type of partitive relation that was presented in Section 7.1.2.

7.2 Terminological relations

As was said at the beginning of this chapter, ***terminological relations*** include relations between terms viewed as lexical units and the meaning they convey. According to this perspective, relations allow us to circumscribe the meaning of a term.

63. It should also be mentioned that some of these relations are opposites: for example, <object-operation> is the opposite of <process-instrument>.

Figure 7.5 shows some relations shared by *habitat* with other terms which were gathered from a corpus of texts on endangered species. Terminological relations can be extremely diversified and complex and link two terms or a larger set of terms. A quick look at the relations in Figure 7.5 for *habitat* can already give an idea of this variety.

Figure 7.5 also shows that some relations considered in lexicon-based approaches differ from those considered in knowledge-based perspectives and that some phenomena are handled differently.[64] For instance, a set of terms with "related meaning" are identified (*environment, range, site, territory*), but none of the terms correspond to a generic or an exact synonym. Additionally, a series of terms denoting activities are linked to *habitat* (*alter, degrade, inhabit, introduce, occupy*) and are expressed as verbal collocates of the term. This section will explain these differences and many more.

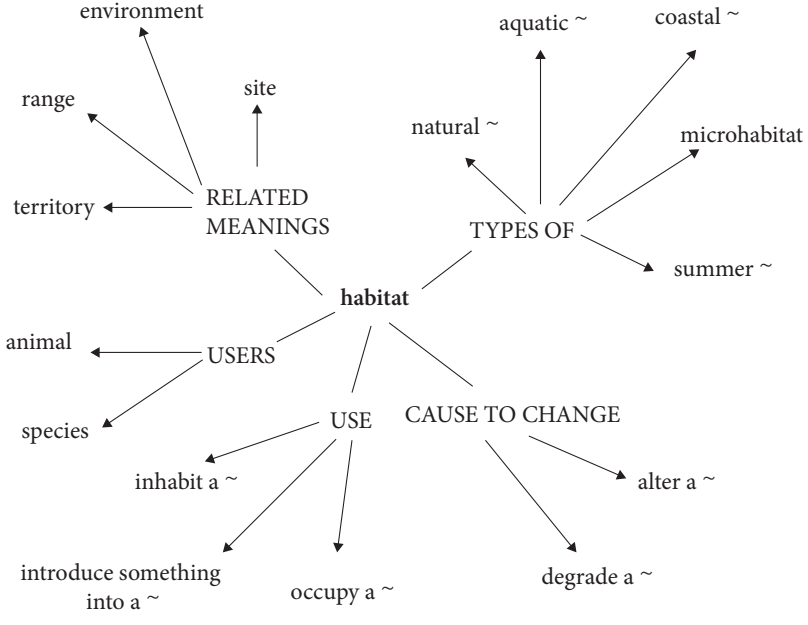


Figure 7.5 Terminological relations between *habitat* and other terms

This being said, some relations considered in both perspectives bear some similarities. For instance, taxonomic relations are also considered in lexicon-driven

64. We can compare the relations that appear in Figure 7.5 with the list reproduced from Sager (1990) in Table 7.2.

approaches. Synonymy and opposition are taken into account, however, their characterization is much more fine-grained in lexicon-based approaches than in knowledge-driven ones. Similarities and differences will be mentioned wherever adequate.

7.2.1 Paradigmatic versus syntagmatic relations

It is customary to classify lexical relations and, by extension, the terminological relations examined in this section, into two categories: paradigmatic and syntagmatic.

Paradigmatic relations, sometimes presented as *vertical relations*, include links between terms within the lexicon. Prototypical paradigmatic relations, such as synonymy, hypernymy, antonymy, connect lexical units or terms that belong to the same part of speech and have the same syntactic distribution, i.e. can fulfill the same syntactic functions. But they also include less classic relations such as those that hold between a verb and a noun that convey the same meaning (e.g. *deforest* ↔ *deforestation*), a verb and a noun that expresses the typical agent (*emit* ↔ *emitter*), and many others. Considering Figure 7.5, the relations between *microhabitat*, *site* and *territory* with *habitat* are paradigmatic.

Syntagmatic relations, sometimes described as *horizontal relations*, include links between terms and other lexical units in a sentence. Syntagmatic relations appear between lexical units or terms that often belong to different parts of speech and that fulfill distinct syntactic functions. In Figure 7.5, the relations between *occupy*, *inhabit* and *summer* with *habitat* are syntagmatic. In this chapter, we consider a subset of syntagmatic relations, i.e. the ones that appear between the components of collocations.

7.2.2 Paradigmatic relations

The three following subsections describe the main families of paradigmatic relations: hypernymy/hyponymy, synonymy, and antonymy.[65] These relations can be

---

65. The status given to relations between LUs, one denoting a whole and another, a part, vary in lexical semantics. For some authors, this set of relations is a central paradigmatic relation: Cruse (1986) devotes a chapter to the subject and WordNet (2017) defines it at a fundamental hierarchical relation for nouns along with hypernymy and hyponymy. For others, it remains much more conceptual in nature than hypernymy and hyponymy, synonymy and antonymy. I consider that it should be taken into consideration in terminological descriptions. I will not devote a new section to it since much of what has been said in Section 7.1.2 still applies when considering terms rather than concepts. However, it is worth mentioning that a specific terminology is used when referring to this set of relations from a lexicon-driven perspective. The LU or term that denotes the part is called a ***meronym***; the LU or term that denotes the whole is sometimes called a ***holonym***.

defined for lexical units or terms that belong to the same part of speech. Other paradigmatic relations hold between lexical units or terms that belong to different parts of speech are presented in Section 7.2.2.4.

7.2.2.1 *Hypernymy and hyponymy*

Hypernymy and ***hyponymy*** are relations of inclusion that connect a more general term, the ***hypernym***, to a more specific term, the ***hyponym***. These relations are closely related to those examined under the general label *taxonomy* and much of what is said in this section will be reminiscent of the content of Section 7.1.1. Knowledge-based approaches refer to hypernyms as *generic concepts* and to hyponyms as *specific concepts*.[66]

As are taxonomic relations in knowledge-based approaches, hypernymy and hyponymy are considered to be fundamental relations for understanding the structure of the lexicon. They are based on the observation that the meanings of terms are included in that of others.

A ***hypernym*** and a ***hyponym*** share most of their semantic content. More specifically, the meaning of the hypernym is included in that of the hyponym, as shown in Figure 7.6. The hyponym has one or a few additional ***semantic components***. *Tool* possesses three semantic components: 'device', 'held in the hand', and 'used for a specific function'; while *hammer* possesses these three components in addition to three new ones: 'with solid head' and 'with handle'. The function is also more precise: 'used for pounding'.

In order to validate the relation between a hypernym and a hyponym, the following contextual tests can be used (Cruse 1986: 137):

An X is a kind of/type of Y

In this test, X stands for the hyponym and Y, for the hypernym. We can apply this test to *hammer* and *tool*.

A hammer is a kind of tool.

However, this test fails when we apply it to other pairs of terms.

?A hammer is a kind of animal.
?A hammer is a kind of screwdriver.
?A handle is a kind of tool.

The test fails in the first sentence because *animal* is not a valid hypernym for *hammer*. It fails in the second because *hammer* and *screwdriver* are ***co-hyponyms***.

66. Of course, as mentioned earlier, *hypernymy* and *hyponymy* can also be used in knowledge-driven approaches, but the distinction is maintained here. It is worth mentioning at this point that the lexical nature of relations is emphasized by the use of terms with the suffix *-nym*.

Finally, it fails in the third sentence since *handle* denotes a part of a tool and not a type of tool.

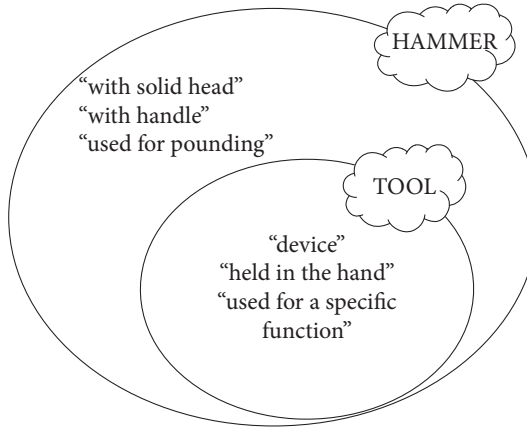


Figure 7.6 Semantic components shared by *hammer* and *tool*

As in a taxonomy, the ***relation*** between hypernyms and hyponyms is ***hierarchical*** and ***asymmetric***. When multiple levels are identified, the fundamental property of ***transitivity*** is maintained. However, some lexical phenomena can disturb the establishment of straightforward hierarchies.

First, a lexical unit or a term might be connected to more than one hypernym depending on the semantic component that is taken into account. An obvious hypernym for the LU *cat* is *feline*. But we could also consider the cat to be a kind of pet. Of course, if we are compiling a terminological resource in any field connected directly or indirectly to biology, *feline* would be the adequate hypernym. However, there might be cases where *pet* would be a better choice, for example, if a terminological resource is concerned with different kinds of stores or the management or a city.

Secondly, when relations are established strictly on the basis of attested LUs or terms, some gaps may appear in a hierarchy. In some cases, it might not always be possible to find names for useful sublevels. For instance, looking at the terms in Figure 7.7, which LUs or terms could distinguish vehicles that move in the air from those that move on the ground; others that are used by personal drivers, from those that must be paid for, etc.?[67]

67. We will see in Section 8.1.2 that knowledge-driven approaches can solve this problem by defining facets to a hierarchy to assist classification efforts and organize concepts more consistently.

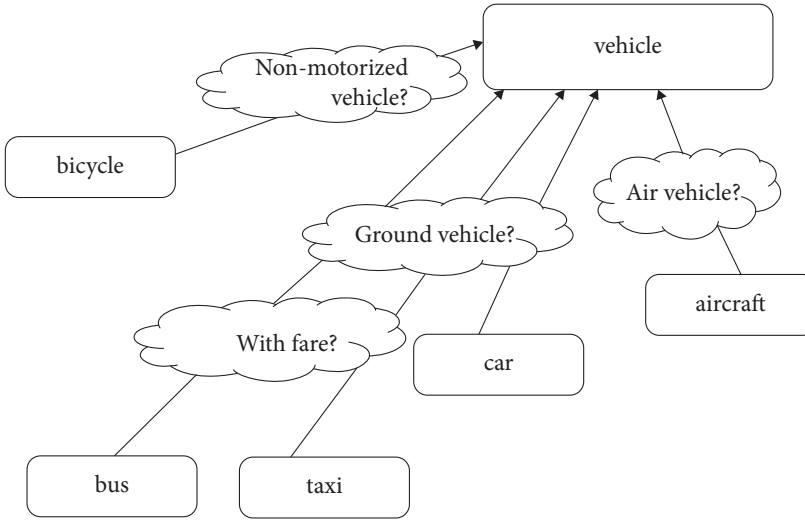


Figure 7.7 Types of vehicles and lexical gaps

It is customary to consider that hypernymic and hyponymic relations apply to lexical units that belong to the part of speech of nouns. Intuitively, inclusion seems to be much more easily applicable to LUs that denote entities. However, inclusion can involve other parts of speech. For instance, the meaning of the verb *move* is included in that of the verbs *walk*, *run*, *cruise* and *swim*. In WordNet (2018), units such as *walk*, *run*, *cruise* and *swim* are defined as ***troponyms*** of *move* and they can be understood as follows: 'move in a certain way'. Similarly, stating that *a scarlet flower entails a red flower* (Cruse 1986: 89) shows that the meaning of adjectives can also be included in that of others.

Although there are some similarities with strict hypernymy, it seems that the extent to which hierarchical levels can be added is much more limited for verbs or adjectives. LUs that express activities or properties (verbs and adjectives) do not lend themselves naturally to the same structuring principles as nouns that denote entities. Consider the example with nouns denoting tools. It would be fairly easy to imagine different levels of a hierarchy in which *tool* would appear (as a hypernym, with *hammer* and *screwdriver*, or as a hyponym, with *instrument*). If we try to do the same with terms that denote different uses of tools, we could probably come up with some verbs (such as *to pound*, *to saw*), maybe with a potential hypernym (*use*?), but we would soon reach a limit beyond which adding levels would be meaningless. In fact, even *use* is very general and not especially informative when connected to verbs like *pound* and *saw*.

7.2.2.2 *Synonymy*[68]

Synonymy is a ***symmetric relation*** between terms or lexical units that have the same meaning or very close meanings. As was mentioned above, knowledge-based approaches to terminology also take into account this relation, but focus on ***exact synonymy***. Lexicon-driven approaches consider other forms of synonymy in addition to exact synonymy.

From the point of view of terminological relations, there is exact synonymy when two terms share all their semantic components. In a pair of ***exact synonyms***, member 1 can replace member 2 in all the sentences where member 2 appears. Conversely, member 2 should be a valid candidate to replace member 1 without affecting the meaning of sentences in which member 1 is used. Furthermore, everything that characterizes one member of the pair is also valid for the other.

Consider *hydropower* ↔ *hydroelectricity* in the field of renewable energy. First, *hydroelectricity* can be used instead of *hydropower* in the following three contexts without altering the general meaning of the statement.

> *Hydropower uses a (mostly) renewable source, (water), and it long has been a favorite electricity source for many countries with large water flows.*
> *(Bjork et al. 2011)*
>
> *Hydroelectricity uses a (mostly) renewable source …*
> *Hydropower holds advantages over other forms of energy – conventional and renewable …*
> *(World Alliance for the Decentralised Energy Association (WADE Thai) 2013)*
>
> *Hydroelectricity holds advantages over other forms of energy …*
> *Around 16 percent of the electricity generated worldwide is produced using Hydropower, according to data from 2007.*
> *(German Energy Association (DENA) 2010)*
>
> *… is produced using Hydroelectricity*

Similarly, *hydropower* can replace *hydroelectricity* in the following sentences without changing their meaning.

> *Waterpower, also called Hydroelectricity, is a renewable form of electricity generation that harnesses the energy produced from the movement of falling or flowing water.* *(Ontario Ministry of Energy and Infrastructure 2010)*
>
> *Waterpower, also called Hydropower*

---

68. For this section, I am grateful to Marjan Alipour who reviewed and suggested criteria for distinguishing synonyms and near-synonyms and classify them more adequately in our environmental resource, the DiCoEnviro.

In 1998, HYDROELECTRICITY provided approximately 21.6% of the worldwide electricity capacity and 18.8% of the worldwide generation of electricity.
… HYDROPOWER provided approximately 21.6% of the worldwide electricity capacity
They may also be used to generate HYDROELECTRICITY.
They may also be used to generate HYDROPOWER

Furthermore, both terms share the same set of terminological relations. They share another synonym (*waterpower*), can be related to the same more general terms or expressions (*source*, *forms of energy*), have the same collocates (*generate* ~, *provide* ~). Finally, they can be explained with the same definition: 'form of energy produced by the power or water'.

Lexical semanticists and lexicographers, when considering the general lexicon, usually state that exact synonyms are extremely rare. It is indeed difficult to identify two lexical units that comply with the conditions mentioned in the previous paragraphs. In contrast, terminologists frequently come across this kind of synonymy.

However, another common situation occurs where two terms share many semantic components but not all of them. In this case, the replacement of a term by the other can be possible in a first set of sentences, but not in others. For instance, *habitat* and *territory* shares many semantic components: (a) they both designate specific kinds of locations; (b) are used by species; and (c) are found in larger areas. Can we assert based on these share semantic components that we are dealing with exact synonyms the same way we did with *hydropower* and *hydroelectricity*?

If we try to replace *habitat* with *territory* in the following sentences, the general meaning of the sentences remains unchanged.

… the HABITATS of many species will move poleward or upward from their current locations. (Intergovernmental Panel on Climate Change (IPCC) 2006c)
the TERRITORY of many species …
protecting plants and animals, and their HABITATS. (EUROPA 2007)
protecting plants and animals and their TERRITORIES
Deterioration of marine HABITATS on the other side of the world affects our food supplies. (EUROPA 2007)
Deterioration of marine TERRITORIES

So far, so good. However, if we try to use *habitat* in sentences in which *territory* appears, the replacement produces odd sentences.

Broods generally remain on the nesting TERRITORY. (Campbell 1995)
?nesting HABITAT
Protection of breeding TERRITORIES and nesting birds from human disturbance is also a priority. (New Mexico Department of Game and Fish 2017)
?breeding HABITAT

This time, the replacement is impossible since *territory* is defined as something used for a specific purpose. Although a *habitat* is also a location where species carry out different activities (live, reproduce, grow, eat); the component 'for a specific purpose' is not a core component of its meaning as it is for *territory*. Hence, we are no longer dealing with exact synonymy. However, since both terms share many semantic components and that a replacement is possible in many (but not all) sentences, the relation between them is one of ***near-synonymy***.

7.2.2.3 *Antonymy and other opposites*[69]

Section 7.1.4 pointed out that knowledge-driven approaches somehow take incompatibility for granted. However, although some forms of opposition can be found in classification systems – for example, 'input device' and 'output device' are defined as opposing categories and can be used to classify more specific concepts –, little attention was given to antonymy *per se*. This can be partly explained by the fact that prototypical antonymy occurs between adjectives (e.g. *compatible* ↔ *incompatible*, *light* ↔ *heavy*) and slightly less prototypical forms of opposition appear between verbs (*install* ↔ *uninstall*, *warm* ↔ *cool*) and not between nouns that denote entities.

In contrast, lexical semantics often regards antonymy as the relation speakers perceive most naturally, and the topic has been addressed in a wealth of analyses (Cruse 1986; Murphy 2003, among others). This section will show that antonymy and other types of opposition can also appear between terms.

Basically, ***antonymy*** links two terms (or more generally lexical units) that share most of their semantic components. This is a feature that was already mentioned for hypernymy/hyponymy and near-synonymy. However, for terms to be ***antonyms*** (or ***opposites***), at least one of their semantic components must introduce an ***opposition***. Below are pairs of terms that are used in the fields of computing and the environment. Not only do both terms share many semantic components, they have the same argument structure. However, one important component sets them apart.

> *compatible* ↔ *incompatible*
> Both adjectives denote the property of computer hardware or software according to which it should be able to work with other hardware or software components.
>
> > *Compatible* (**X** is compatible with **Y**) means that the computer hardware or software is able to work with other hardware or software components.

69. The content of this section is based on an article published in *Terminology* (Gagné & L'Homme 2016). I am grateful to Anne-Marie Gagné for our numerous discussions on the topic of opposition.

Incompatible (**X** is incompatible with **Y**) means that the computer hardware or software is not able to work with other hardware or software components.

download ↔ *upload*
Both verbs denote activities in which a file is moved from one location to another.

Download (**X** downloads **Y** from **Z** to **W**) means that the file is moved from a remote location to a local computer.
Upload (**X** uploads **Y** from **Z** to **W**) means that the file is moved from a local computer to a remote location.

warming ↔ *cooling*
Both nouns denote events in which the temperature of something (the climate, the ocean) changes.

Warming (warming of X) means that the temperature is rising.
Cooling (cooling of X) means that the temperature is decreasing.

Even if all the terms above are antonyms, the specific opposition is triggered by different factors. *Incompatible* versus *compatible* can be roughly characterized as 'not compatible' or 'something that is not compatible is necessarily incompatible'. This first characterization does not apply to the *download* ↔ *upload* pair (not downloading something does not entail that we are necessarily uploading it) or to *warming* and *cooling* (a 'cooling' situation does not include all 'non-warming' ones). *Download* and *upload* as well as *warming* and *cooling* are still opposites, but in a different way. Opposite relations are diverse and cannot all be characterized with the same set of criteria.

Interestingly, some terms have more than one opposite. For instance, *polluting* can be opposed to *non-polluting*, to *green* and *clean*, and even to *polluted*. However, the opposition between *polluting* and these other terms depends on different factors.

– **One or the other**

A first type of opposition, and the most basic and straightforward one, is found between two terms (often adjectives) where each one completely fills one of the two parts of a semantic domain. In this form of opposition, something has to be one or the other; nothing can be both. Terms that are opposed this way are called ***complementary antonyms***.

In a pair of complementary antonyms, the denial of one member automatically entails the assertion of the other. We can validate this with the pairs *soluble* ↔

insoluble and *organic* ↔ *inorganic* as shown below. This is also the sort of opposition that appears between *compatible* and *incompatible* and between *polluting* and *non-polluting*.

> *This compound is not* SOLUBLE, *then it is necessarily* INSOLUBLE, *that organism is* SOLUBLE, *then it necessarily is not* INSOLUBLE.
> *This matter is not* ORGANIC; *then it is necessarily* INORGANIC; *this matter is* ORGANIC; *then it is not* INORGANIC.

– **Opposing values on a scale**

A second type of opposition involves terms that need to be considered with respect to a scale.

In order to understand this kind opposition, we must picture a variable property or phenomenon, such as height, length or weight (Cruse 1986). Consider, for instance, the variable property of temperature. Temperature can have a wide range of values as meteorologists remind us every day. Among the possible values that temperature can have are those that can be considered to be high or low (Figure 7.8). In English, the opposing values of temperature are expressed with the adjectives *warm* (for the high value) and *cool* (for the low value).[70] The high and low values are regarded as opposites with respect to a central point where temperature would be neither high nor low. The terms that express these values are called ***gradable antonyms***.

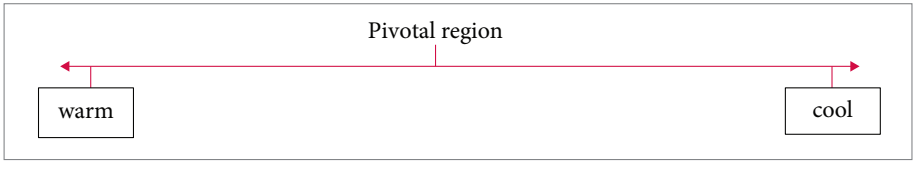


Figure 7.8 A gradable pair situated on the scale of temperature (according to Gagné & L'Homme 2016)

Gradability often characterizes opposition between adjectives since they denote properties. In a pair of gradable antonyms, both members of a pair cannot be valid at the same time. This can be confirmed with the following test: if it is X, then it is

70. It should be specified that, in this example, temperature is considered from the point of view of climate change. This is why *warm* and *cool* can be perceived as gradable antonyms. In other fields or in general language, the usual antonyms proposed with respect to temperature are *hot* and *cold*.

not Y; or if it is Y, then it is not X. The test is applied below to the adjectives *humid* and *dry* and to the pair *cool* ↔ *warm*.

> *If the air is* DRY*, it is not* HUMID *and if the air is* HUMID*, it is not* DRY*.*
> **The air is* DRY *and* HUMID

> *If the temperature is* COOL*, then it is not* WARM *and if the temperature is* WARM*, then it is not* COOL*.*
> **The temperature is simultaneously* WARM *and* COOL*.*

Gradable antonyms differ from complementary antonyms in the sense that if X is not Y, it does necessarily entail that X is Y. It could be something else on the scale. In some situations, neither member is valid. Finally, a situation can be characterized by 'more or less'; situation A is X, but it is more Y than situation B.

> *The air is* HUMID*, but it is dryer than yesterday's.*
> *The air is comfortable; it is neither dry nor* HUMID*.*

> *The temperature is* WARM*, but it is cooler than yesterday's.*

Interestingly, a pair of lexical units may have been characterized as gradable antonyms in general language, but defined as complementary in a specialized domain. This situation occurs with *light* and *heavy* in the field of electric vehicles.

> *In contrast to the self balancing devices,* LIGHT *electric vehicles are emerging which are similar in design to a bicycle, but require no pedaling.* *(Rose 2011)*
> *It will have to meet the needs of both the* HEAVY *vehicle industry (vehicles used in public transit, for example) and the light vehicle industry.*
> *(Gouvernement du Québec 2012)*

In this particular field of knowledge, *light* and *heavy* each fill one of the two parts of a semantic domain. The sense of scale that applies to gradable antonyms is lost since a vehicle has to be classified as either light or heavy. This might be another effect of meaning modulations that were mentioned in Section 5.2.5.

– **Opposite directions**

Terms that denote activities – verbs and nouns – can be opposed in ways that differ from the types of opposition that were characterized so far. These opposites, called ***reversives***, appear in two different situations.

In one situation, the terms represent a different perspective on a situation; they express a change of direction between two absolute states. The initial state of the first member corresponds to the final state of the second member and vice versa. The examples below with *afforestation* and *deforestation* illustrate this first case of reversiveness. The initial and final states 'with trees' and 'without trees' are considered to be absolute because they mean something by themselves.

… DEFORESTATION of smaller areas within a unit may not take the canopy cover of the unit below the forest definition threshold.
(Intergovernmental Panel on Climate Change (IPCC) 2006e)
AFFORESTATION of degraded forests and wastelands are the most attractive opportunities. (Intergovernmental Panel on Climate Change (IPCC) 2006b)

Deforestation denotes an activity whereby someone removes trees from an area. This area corresponds to a **Source** in terms of semantic roles (see Section 6.5). In contrast, *afforestation*, denotes the activity in which someone places trees in the area. In this case, the area corresponds to a **Destination**.

The same situation applies to the verbs *install* and *uninstall*. To *install* a program consists in placing a program on a computer or more directly on a storage device; to *uninstall* it consists in removing the program from where it had been previously installed. Again, the place where the program is installed is a **Destination**; but is a **Source** when considered with regard to *uninstall*.

The operating system is INSTALLED on your hard disk. (LINUX Directory 2004)
This is the applet you should use to UNINSTALL programs that you no longer want. (Computers for Beginners 2004)

A second group of ***reversives*** concerns two opposite movements between relative states, meaning that these states only make sense when referring to each other. This is the kind of opposition that can be observed between the verbs *melt* and *freeze*.

… experts predict smaller mo untain glaciers could MELT. (UNESCO 2006)
River and lake ice will break up earlier and FREEZE later.
(Intergovernmental Panel on Climate Change (IPCC) 2006a)

This second situation, in contrast with the first one, does not entail a change in the semantic roles of arguments. The initial and final states are relative to each other. In the case of the *melt* ↔ *freeze* pair, something might start melting but still be perceived as frozen.

Two other relations described below also introduce a type of opposition, but they are not defined as antonymy *per se*.

– **Different perspectives on the same situation**

This first case examined is that of ***converse terms***. Here, the sense of opposition arises from a semantic symmetry resulting from the permutation of arguments. Consider the examples below:

The mouse precedes the cat.
The cat follows the mouse.

These two sentences are logically equivalent: they mirror each other. This perceived equivalence arises from a change of position between arguments: *mouse*, the first argument of *precede,* becomes the second argument of *follow. Cat* also changes status: it is the second argument of *precede*, but the first argument of *follow*.

This relation is also shared by *send* and *receive*. The sender is the first argument of *send* (**X(sender)** sends **Y** to **Z(receiver)** by **W**), while it becomes the third argument of *receive* (**X(receiver)** receives **Y** from **Z(sender)** by **W**).

> *This means that when someone* SENDS *you a DOC file by email, it may be more than just a data file.* *(Hruska 2001)*
> *If you* RECEIVE *an email about an auction that you have not participated in, and there are instructions to visit a web site to cancel the order, don't do it.*
> *(An Internet Guide to Newcomers to the World Wide Web 2004)*

Converseness can also be observed between quasi-predicative terms as in the examples below with *predator* and *prey. Predator* is defined as **X** is a predator of **Y;** **X** corresponds to the argument realizing the predator; and **Y** to the prey. P*rey*, on the other hand, is defined as **X** is a prey of **Y**. In this case, **X** is the argument realizing the prey and **Y,** the predator.

> *Predators of musk deer include the lynx, wolf, fox, and yellow-throated marten (mammal related to the weasel and the mink).* *(Benson & Nagel 2004)*
> *Eels appear to be the preferred* PREY, *although other fish, tadpoles, and salamanders are also occasionally eaten.* *(Mississippi Museum of Natural Science 2014)*

Converse terms do not always introduce opposition or the opposition is so subtle that it might be difficult to perceive as such. Consider the following example with *run* and *propel*.

> *Nearly every automaker is announcing vehicles that can plug in and* RUN *on electricity.* *(Axsen et al. 2011)*
> *Hybrid and PHEVs are able to blend electric and engine power to* PROPEL *the vehicle and therefore require less total onboard power than the E-REV.* *(Tate et al. 2008)*

Technically, these terms are converse, since their arguments are permuted. *Run* is defined as **X (a vehicle)** runs on **Y (an energy source)**, while *propel* is defined as **X (an energy source)** propels **Y (a vehicle)**. However, in this case, there is not true opposition.

– **Opposition based on conventions**

A last group of terms are perceived as opposites, but this opposition cannot be characterized in a way that applies to the categories mentioned so far. It involves terms that denote entities and is often based on conventions rather than on precise

semantic components. Often more than one semantic component opposes the two terms, which are called ***contrastives***.

This is the form of opposition that can be observed between *software* and *hardware*. It is customary in the field of computing to contrast the two terms. *Hardware* denotes all the physical equipment (the computer and all the physical entities that surround it); *software*, on the other hand, denotes all the programs, applications, and operating system used on the computer.

This relation also applies to *continent* and *ocean* in the environment. These two entities both cover different parts of the Earth, and it is customary to oppose them. *Land* and *sea* as well as *fauna* and *flora* are also defined as contrastives in the same domain.

This last category of opposites is probably the most compatible with the types of opposites considered in knowledge-driven approaches (see Section 7.1.4).

7.2.2.4 *Paradigmatic relations across different parts of speech*

The relations examined so far – hypernymy and hyponymy, synonymy and antonymy – hold between terms or lexical units that belong to the same part of speech. Less prototypical ***paradigmatic relations*** link terms or lexical units that can belong to different parts of speech. Consider the following sentences that contain terms that are linked to the intransitive verb *erode*:

> *Many of the freshwater coastal marshes would become salt marshes, and soft coastlines would* ERODE *more rapidly.* *(Environment Canada 2007)*
> *Many dryland areas face increasingly low and erratic rainfalls, coupled with soil* EROSION *by wind and the drying up of water resources through increased regional temperatures.* *(Müller & Buchdahl 2000)*
> *Rising sea levels and* ERODING *coastlines are already a major threat to exterior provinces ….* *(Environment Canada 2007)*
> *Soils with high organic matter content are less* ERODIBLE *than those with low organic matter content.* *(PANACEA 2015)*
> *Many exposed subsurface soils on* ERODED *sites tend to be more* ERODIBLE *than the original soils were, because of their poorer structure and lower organic matter.* *(PANACEA 2015)*

This first set of examples show that some terms convey the same meaning: *erode* and *erosion* both denote a process that affects soil. As far as meaning is concerned, there is no difference between the noun and the verb; only the part of speech is affected and consequently so is the syntactic distribution of terms. Other terms are closely related to *erode* but include additional semantic components: *eroded*, *eroding* and *erodible* apply to entities that undergo the erosion, but in slightly different ways.

– **Different parts of speech/same meaning**

Erode and *erosion* have the same meaning but this meaning is realized as a verb in one case ('soil gradually degrades and is worn away') and as a noun in the other ('process whereby soil gradually degrades and is worn away').

This situation is quite common in specialized domains: many verbs and nouns carry the same meaning: as the pairs *pollute* ↔ *pollution, connect* ↔ *connection*, *electrify* ↔ *electrification, process* ↔ *processing* shown in the examples below:

> *The salt water can penetrate aquifers – not to mention wells – and* POLLUTE *the water.* *(PANACEA 2015)*
> *In Canada,* POLLUTION *of surface water by groundwater is probably at least as serious as the contamination of groundwater supplies.* *(Environment Canada 2007)*

> *Anyone can* CONNECT *to either of these networks ….* *(Curtin 2004)*
> *As an end user, your only concern is that the* CONNECTION *is good, but for a network engineer, this can mean several different types of technologies.* *(Hruska 2001)*

> *The government will* ELECTRIFY *its ministry and agency fleets.* *(Gouvernement du Québec 2012)*
> *The focus of the European Commission and other EU institutions on electric mobility … has only recently began to shift towards the* ELECTRIFICATION *of bus systems in Europe.* *(Wołek & Wyszomirski 2013)*

> *A computer is an equipment used to analyze and* PROCESS *data and display the results on any medium it is capable of writing on.* *(Castro 2001)*
> *This* PROCESSING *of incoming data is usually handled by a script or program written in Perl or another language that manipulates text, files, and information.* *(A Beginner's Guide to HTML 2001)*

Nouns and adjectives can also share the same meaning. This is shown with *compatible* and *compatibility* and between *abundant* and *abundance* below.

> *Without* COMPATIBLE *applications at both ends, the data sent doesn't end up anywhere….* *(Linux Directory 2004)*
> *… check your HTML for browser* COMPATIBILITY. *(Introduction to Computers 2004)*

> *… the species can be locally* ABUNDANT *in suitable habitat.* *(Lindenmayer et al. 2011)*
> *… the total range and* ABUNDANCE *of the species must be considered.* *(Stohlgren & Sunil 2013)*

Terms that share the same meaning also have the same argument structure (abundant **X**; abundance of **X; X** is compatible with **Y**; compatibility of **X** with **Y; X**

erodes; erosion of **X**). Even if the order or arguments changes, their number is the same and their linguistic realizations are semantically related. In addition, if they were labeled (with semantic roles or another system), the labels would be identical (e.g. **Patient** erodes; erosion of **Patient**).

– **Different parts of speech/closely related meanings**

Erode is also semantically related to at least three different adjectives, i.e. *eroding*, *eroded*, and *erodible*. However, the meaning of each adjective differs slightly from that of the verb, and the meanings of the three adjectives differ from one another. *Eroding* means that something is currently undergoing the process expressed by *erode; eroded* means that something already underwent the process; *erodible* means that something may undergo the process expressed by the verb. The same patterns can apply to other series of terms (e.g. *pollute, polluted, polluting; compost, compostable*).

Again, closely related terms share some of their arguments (**X** erodes; eroding **X**; erodible **X**; eroded **X** – **X** pollutes **Y** with **Z**; polluted **Y**; polluting **X**). The linguistic realizations of the arguments of *erode*, *eroding*, erodible, and *eroded* are of a similar nature (*beach*, *soil, coast, coastline*). In the same way, the terms that would instantiate **X** in sentences with *pollute* and *polluting* would be units such as *particle*, *substance*, *vehicle*; those that would instantiate **Y** in sentence with *pollute* and *polluted* would be units like *air, lake*, *pond*, *soil*, etc.

– **A special kind of adjective**

Relational adjectives are quite recurrent in specialized domains and have drawn the attention of terminologists since they are closely related to noun terms. Although more commonly used in Romance languages, they can still be found in English as shown below.

> *For cold countries such as Canada,* CLIMATE *change can indeed provide some significant benefit ….* *(Environment Canada 2007)*
> CLIMATIC change is a stimulus to the migration of both plants and animals. (Environment Canada 2007)

> The small concentrations of greenhouse gases within the ATMOSPHERE that cause this effect …. (Environment Canada 2007)
> ATMOSPHERIC aerosols influence the transfer of energy in the atmosphere in two ways …. (Müller & Buchdahl 2000)

Relational adjectives share much of their semantic content with the noun to which they are related. However, they behave differently from other adjectives such as *abundant* that was mentioned above.

– First, although both types of adjectives can be used as modifiers, the relational adjective cannot be used in predicative position:

abundant species
climatic change

the species is abundant
**the change is climatic*

– Secondly, most adjectives can be modified by an adverb of degree; this is not possible for relational adjectives:

the species is very abundant; a very abundant species
**a very climatic change*

Both types of adjectives also differ semantically. Adjectives such as *abundant* express a quality, a property of the object or activity denoted by the modified noun. Furthermore, as was said above, they have the same argument structure as the corresponding noun. In contrast, relational adjectives express the argument of the noun they modify syntactically.

X is abundant ; abundance of X
climatic change → change in climate

– **Arguments and circumstantials**

Another set of regular relations between terms that can belong to different parts of speech concern predicative terms and their arguments or circumstantials. Consider the following examples:

The other is as a PROGRAMMER*, where you are writing shells scripts or programs, which can be written in a variety of ways (Tsariounov 2004).*

… you may EDIT *the current page in the HTML* EDITOR *of your choice.*
(Cohen 2004)

Do you believe that we are destroying the earth by drilling and mining for oil and coal and then POLLUTING *the atmosphere with toxic emissions?*
(PANACEA 2015)
Ozone is a greenhouse gas. It also shields the surface of the Earth from harmful ultraviolet (UV) radiation, and is a common air POLLUTANT*.*
(Intergovernmental Panel on Climate Change (IPCC) 2006b)

If you use Outlook, disable HTML messages to prevent e-mails that are FORMATTED *as Web pages from being displayed. (MISSING SOURCE)*
As with every multimedia FORMAT *there are a number of types.*
(Introduction to Computers 2004)

Programmer is the name used for the typical **Agent** of the verb *program; editor* is the **Instrument** used for *editing*; a *pollutant* is the name of the typical **Material** that causes the process denoted by *pollute*; and, finally, *format* is the **Result** of *formatting*. These relations and other ones that link an argument or a circumstantial to a predicative unit are quite common in specialized domains.

Again, the relation between the terms is visible in their argument structures:

X programs **Y** with **Z: X** is a programmer of **Y**;
X edits **Y** with **Z**; editor used by **X** to act on **Y.**

– **Formal and semantic relations**

In specialized domains, ***morphologically related terms*** often share a semantic relation. A shared ***stem*** or a shared ***morpheme*** can be indicative of a shared semantic component as shown in the examples below:

Shared morphemes
-ist (Environment: someone who specializes in)
biologist (source term: *biology*)
climatologist (source term: *climatology*)
ecologist (source term: *ecology*)

bio- (Environment: organic matter)
biodiesel (source term: *diesel*)
bioethanol (source term: *ethanol*)
biofuel (source term: *fuel*)
biogas (source term: *gas*)

-er/or (Computing: a person who carries out a specific task)
developer (source term: *develop*)
programmer (source term: *program*)
user (source term: *use*)

Shared stem
pollut- (undesirable and harmful presence in a location)
pollution
depollute
polluter
pollutant
polluting
polluter

It should be kept in mind, however, that other terms can share the same semantic relations without having any formal relations with one another. For instance, persons who carry out specific tasks can be designated by terms that do not carry the

suffix *-er* (e.g. *computer scientist, expert*). Additionally, a given morpheme can be associated with more than one meaning. For instance, *-ist* is used to create terms that denote experts in a given field (*biologist*, *climatologist*, *ecologist*). However, the term *environmentalist* denotes a person who has convictions about protecting the environment (and this person is not necessarily an expert in any of the disciplines studying the environment). Similarly, the suffix *-er/or* can be used in computing to designate specific kinds of instruments (e.g. *compiler, browser, editor*) and not always someone associated with a given task.

7.2.3 Syntagmatic relations

An important set of terminological relations differ from the ones presented so far in the sense that they concern the way terms are combined with other terms or lexical units in sentences. This second set of ***relations*** are ***syntagmatic***.

Terms combine with other lexical units or terms according to two principles. First, combinations comply with syntactic rules. Nouns, verbs, adjectives and adverbs can be combined with a limited set of other parts of speech and they fill predefined syntactic functions. For instance, an adjective (or adjective phrase) modifies a noun or is used in a predicative position.

Abundant species can be found in this area.
The species is abundant in this area.

In addition to these syntactic constraints, terms and lexical units can be combined if they share semantic affinities. *Abundant* can be used with *species*, *plant*, and other terms that denote living organisms. However, it is odd to use it with locations, psychological dispositions or names of domains.

Abundant species, abundant plants, abundant vegetation
?*Abundant areas, abundant concerns, abundant ecology*

These two rules are regular and apply to all terms and lexical units. However, other combinations obey more subtle constraints. For instance, in order to express the idea of 'creating' with the term *program*, although in theory, different options are possible (*create, develop, design*), the "best" choice would be *write a program*. To express the same idea of 'creating' with *application*, the most idiomatic choice would be *to develop*. Groups such as *write a program* and *develop an application* are ***collocations***.

Collocations have been characterized in general language as linguistic expressions in which one lexical unit – called the ***base*** – is selected freely by a speaker and the second lexical unit – called the ***collocate*** – must be used with a given

base to express a specific meaning (Mel'čuk 1996; Polguère 2016: 66).[71] Typical examples of collocations are *ask a question* and *heavy smoker. Question* and *bachelor* can be selected freely in a given communicative situation. However, only *ask* can be used to express the idea of 'making a question'; *heavy* must be used with *smoker* to express the idea of 'intensity'. In French, *ask a question* translates into *poser une question* (lit. place a question); *heavy smoker* into *gros fumeur* (lit. large or fat smoker) or *fumeur invétéré* (lit. inveterate smoker).

Collocations are also common in specialized domains (e.g. *write a program*, *browse the Web* in computing and the Internet; *be at risk*, *endangered species* in the environment; *ride a bike* in cycling). As in general language, the constraints that explain the combination of two components of a collocation cannot be fully explained by stating general linguistic rules. Rather, they are unpredictable and comply with usage in a given field of knowledge. This is why terminologists have become increasingly interested in collocations and have started adding them to specialized resources (Binon et al. 2000; Cohen 1986; TERMIUM Plus® 2017, to name a few).

Summary

Relations between concepts (conceptual relations) or terms (terminological relations) are central in terminology. However, depending on the approach, different kinds of relations and even closely connected relations are handled with specific criteria.

Conceptual relations include taxonomies and partitive relations. Knowledge-based approaches usually focus on exact synonymy rather than near-synonymy and define it at a relation between designations. Recent work takes into account a wider range of conceptual relations, such as <cause-effect>, <object-attribute>, some of which are crucial to the understanding of concepts in certain domains.

Terminological relations fall into two separate groups, paradigmatic and syntagmatic. Paradigmatic relations include different kinds of synonymy, hypernymy/hyponymy as well as antonymy and other forms of opposition. These relations hold between terms that belong to the same part of speech. Paradigmatic relations also include various links between terms that belong to different parts of speech. Syntagmatic relations considered by terminologists are called *collocations*.

---

71. This is one of the characterizations of collocations that were proposed and the most compatible one with the approach taken in this book. Others approach collocations as a statistical phenomenon.

This chapter introduced various relations between concepts or terms from the point of view of a single language. The next chapter examines how complex structures can be derived from large sets of relations once they are established.

Further reading

The third chapter of Sager (1990) is a thorough presentation of conceptual relations and different issues related to their definition in terminology. Nuopponen (2014) discusses properties of taxonomies and partitive relations. A classic and often cited reference regarding partitive relations and their different subtypes is Winston et al. (1987). Nuopponen (1994), Barrière (2002) and Marshman (2006) analyze cause-effect relations in specialized domains. Various conceptual relations are described in the resource EcoLexicon (2018; Faber et al. 2016). Exploring the knowledge base gives a flavor of the complexity of conceptual relations that need to be handled in a broad domain such as the environment.

Cruse (1986) is a must read for anyone interested in lexical relations, especially the complex properties of synonymy, hypernymy/hyponymy, meronymy and antonymy. Lexical relations are also discussed at length in Murphy (2003) and Palmer (1976). An interesting cognitive interpretation of relations can be found in Croft and Cruse (2004).

Amsili (2003) and Gagné and L'Homme (2016) deal with antonymy and opposites in terminology. Relational adjectives in specialized domains were studied by Daille (2001) and Maniez (2015). Collocations have attracted interest in terminology since the 80s. Cohen (1986) is the first specialized dictionary listing collocations in the field of the stock exchange.

CHAPTER 8

Discovering structures in specialized domains

Terminology assumes that concepts or terms pertaining to a domain are *connected* through a network of relations. So far, we described individual relations and how they apply to small sets of concepts and terms. In practice, however, terminologists must consider larger and more complex networks of relations that emerge from a special subject field. This chapter explores this matter and explains how different kinds of structures can be derived from the relations examined in Chapter 7.

A structural system gives a broader perspective on specialized knowledge. It furthers our understanding of a domain since it provides an explicit interpretation of how concepts or terms interact. Structures can also be exploited to obtain different views on specific areas of knowledge.

The fundamental distinction between conceptual and terminological relations that was made in Chapter 7 will be maintained in this one since, as can be expected, a different perspective on relations leads to different kinds of structures. Hence, the chapter is divided into two main sections: Section 8.1 examines systems derived from conceptual relations; Section 8.2 describes different methods for developing structures based on terminological relations.

8.1 Structures based on conceptual relations

The structures studied in this section – ***ontologies***, ***thesauri***, ***terminological knowledge bases*** – use the relations described in Section 7.1, especially taxonomic and partitive relations. Other conceptual systems implement a much broader set of relations and will also be mentioned.

Conceptual structures comply with two basic prerequisites. The first one dictates that there must be a clear and formal distinction between concepts and the labels used to express them in language. This distinction is made the General Theory of Terminology (GTT) in which ***concepts*** are clearly separated from ***designations***. This formal separation is necessary in all conceptual systems, even those that are not directly grounded in the principles of the GTT. The examples presented in the following sections label concepts with English units or expressions ('nervous system' in the Foundational Model of Anatomy (Figure 8.2) or 'habitat'

in EcoLexicon (Figure 8.3)). However, it should be kept in mind that all nodes in conceptual structures represent concepts. It is more convenient to use labels that can be understood by users, but some conceptual systems resort to more neutral expressions, such as alphanumeric codes.

Some concepts can be expressed by more than one designation, leading to a situation of ***exact synonymy*** (Section 7.1.3). Many conceptual structures record exact synonyms since it is useful for users to know the different names given to concepts in a language. An example was given in Figure 7.4 with 'mainframe' to show how exact synonyms are attached to a concept. In a conceptual structure, the information on exact synonyms is *appended* in a separate module and should not affect the structure directly. This method also implies that a ***preferred designation*** be chosen among possible synonyms.

The second condition with which conceptual structures must comply is that they must be consensual. Concepts are delineated in a way that is admitted within a sufficiently large community. This second requirement implies there is no overlap between concepts, i.e. concepts are mutually exclusive. In some fields of knowledge, experts spend a lot of time agreeing on definitions for concepts and selecting relevant characteristics to classify them.[72]

8.1.1 Accounting for and representing conceptual relations

Structures can highlight different kinds of ***conceptual relations***. Often, these structures are first organized around taxonomic relations since they are believed to constitute the backbone of knowledge organization (Section 7.1.1). Other relations (partitive, <cause-effect>) can then be added to a first taxonomical organization of concepts.

The simplest structures take into account a single relation. Examples were given in previous chapters: Figure 2.1 illustrates the way the concept 'Vulpes vulpes' is connected to broader classes, such as 'Vulpes' and 'Canidea'; Figure 2.6 shows how the concept 'habitat' is connected to the generic 'environmental system' and the specific concepts 'terrestrial habitat' and 'arbotereal habitat'. These two structures are based on ***taxonomic relations***. They help us view the objects of a specialized domain as small groups that can be further organized into larger groups, or, conversely, as large groups that can be broken down into smaller groups. Although less common, other single-relation structures are based on ***partitive relations***. An example was given in Figure 7.3 with bicycle and some of its parts, i.e. 'wheel', 'frame'.

---

72. More recent knowledge-driven approaches take a more relaxed approach on relations between terms. However, there is always a point where concepts must be clearly distinguished from one another.

The preferred representation of ***hierarchical relations*** is a tree structure where the genericity of generic concepts over subordinate ones can be readily perceived. Figure 8.1 is a textual version of a taxonomy of musical instruments (a graphical version of this structure was given in Figure 7.1).

- musical instruments
  - + woodwind instrument
  - brass instrument
    - trumpet
    - …
  - string instrument
    - violin
    - cello
    - double bass
    - …
  - percussion instrument
    - cymbals
    - harp
    - xylophone
    - …

Figure 8.1 Textual representation of a taxonomy of musical instruments

Some terminological structures combine two kinds of relations. Consider the example in Figure 8.2 taken from the Foundational Model of Anatomy (2018). The concept 'nervous system' is connected through a taxonomic relation to 'organ system' (a generic concept). The same relation links 'nervous system' to 'nervous system of female adult body' and 'nervous system of male adult body' (the latter are specific concepts of 'nervous system'). However, other sets of concepts are connected through partitive relations: the 'autonomic nervous system' is defined as a regional part of the 'nervous system'. The former is further broken down into two parts, i.e. 'central part of autonomic nervous system' and 'peripheral part of autonomic nervous system'.

Of course, besides taxonomic and partitive relations, many other kinds of relations can be found between concepts: <cause-effect>, <entity-function>, <activity-instrument>, etc. (see Table 7.2 for a longer list of alternative relations). In certain domains, some of these relations can be considered equally or even more important than hierarchical relations. Conceptual systems can take into consideration multiple relations. In these types of structures, the same concept is linked to several other concepts, but through different relations. For example, 'bicycle' shares taxonomic relations with 'vehicle' (its generic concept) and 'electric bicycle' (a specific concept). It also holds a partitive relation with 'wheel' and 'handle bar'.

We could add other less prototypical relations to further characterize the position held by 'bicycle' in the domain of cycling. It holds an <object – operation> relation with 'riding'; it also shares an <object-functioning> with 'electrically-propelled'. By adding this new relations, we obtain a deeper understanding of the concept.

Figure 8.2 Taxonomic and partitive relations in the Foundational Model of Anatomy (2018)

Theoretically speaking, being able to access the entire set of relations between all the concepts in a given domain is highly desirable. However, representing them can quickly raise some challenges in practice. In the terminological knowledge base EcoLexicon (2018; Faber et al. 2016), a high number of relations are defined between concepts. The graphical display chosen to visualize them is a ***graph***.

Figure 8.3 shows the relations held by 'habitat' with other concepts. Concepts correspond to nodes with English labels. Relations between two concepts are represented as edges which are labeled according to the direction in which the relation is expressed. In addition to taxonomic relations ('intertidal habitat' <type-of> 'habitat'; 'canopy' <type-of> 'habitat') and relations defined as partitive ('colonization' <located-at> 'habitat'; 'guild' <located-at> 'habitat'), many more relations are taken into consideration: for instance, 'ecology' <studies> 'habitat'. In the online version of the resource, edges that stand for different sets of relations are distinguished by color.

In resources that implement various types of relations between large sets of concepts, the resulting representation can soon become overwhelming. This situation occurs in EcoLexicon, since the resource covers many environment concepts and is still enriched on a regular basis. Different methods were developed to allow users to focus on smaller portions of the conceptual structure. First, users can parametrize the depth of the graph to two or three levels in order to limit the complexity of the structure displayed (the graph in Figure 8.3 has two levels). Another way to visualize a more focused content consists in selecting a set of relations: in

EcoLexicon (2018), relations are distributed in three different families (generic-specific; part-whole and non-hierarchical). Finally, users can navigate through the conceptual structure by selecting any node and generate a new graph in which the selected node will become the central concept. An additional feature also allows users to view portions of the graph that are associated with given topics of the environment. It is presented in the next section.

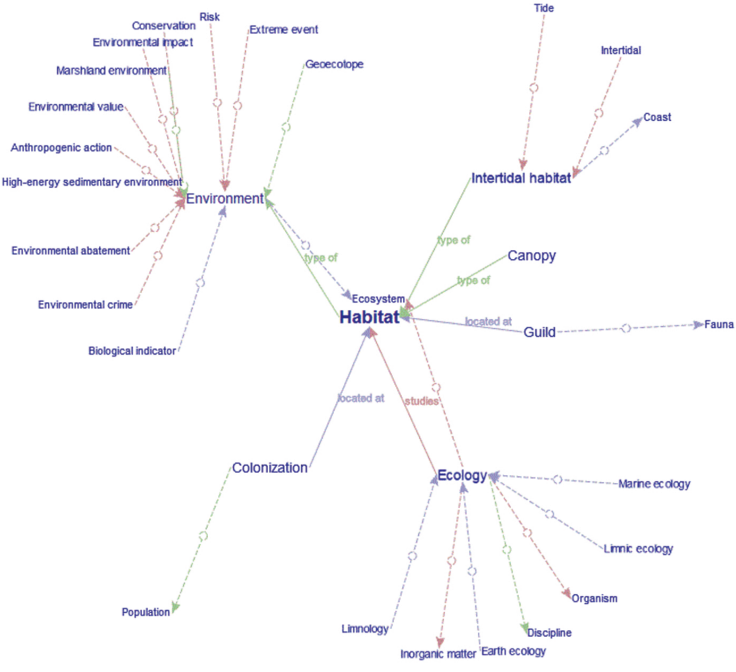


Figure 8.3 Conceptual relations held by 'habitat' and other concepts (EcoLexicon 2018)

8.1.2 Handling relations in conceptual structures

Once a large number of concepts are formally connected through explicit relations, different kinds of operations can be carried out. This section examines some of these operations assuming that they are applied to a conceptual structure in which the backbone is a taxonomy.

– **Passing on characteristics**

Chapter 7 (Section 7.1.1) explained how generic concepts lend all their characteristics to specific concepts in a taxonomy. In a conceptual structure, we can formally establish that there is ***inheritance*** between a generic and a specific concept. For instance, let us postulate that the following characteristic captures the concept 'musical instrument': 'used to produce music'. When 'string instrument' is defined as a specific concept of 'musical instrument', it inherits this characteristic. 'String

instrument' has a few more characteristics with respect to 'musical instrument': i.e. 'has strings', 'produces sounds when strings are vibrating'. If other specific concepts are added, i.e. 'wind instrument', 'bass instrument', they also inherit the 'used to make music' characteristic from 'musical instrument'.

Many taxonomies have more than two levels and the relation between a generic and its immediate subordinate concepts is also valid for the other concepts situated at lower levels. We can exploit this property called ***transitivity*** by further propagating the characteristics assigned to a generic concept to lower levels of the hierarchy. For instance, if 'violin' is defined as a specific concept of 'string instrument', it inherits the characteristics 'has strings', 'produces sounds when strings are vibrating' from 'string instrument'. Furthermore, since 'string instrument' inherits the characteristic 'used to produce music' from 'musical instrument', this characteristic is passed on to 'violin'.

– **Formalizing incompatibility**

Section 7.1.4 mentioned that a taxonomy includes ***incompatible concepts***. Indeed, specific concepts that fall under the same generic, such as 'string instrument' and 'wind instrument', are mutually exclusive. No object can belong to both categories. Incompatible ***concepts*** can be formally declared as ***disjoint*** and the characteristics assigned specifically to one concept cannot apply to the second.

– **Assisting classification**

In taxonomies, categories can be created in order to organize knowledge according to a selected set of characteristics. This was done in the standard animal species taxonomy. Different levels such as Phyllum, Class, Order, and so on are designed to capture the entire set of known animals and result from ample discussion among experts. Their existence is justified for classification efforts.

Section 5.4 showed that sets of concepts may lend themselves to different classifications leading to a phenomenon called ***multidimensionality***. A first form of multidimensionality occurs when different characteristics are used for the classification. In a conceptual structure, this kind of multidimensionality can be managed with ***facets***. Facets account for different ways in which specific concepts are connected to a generic one. The example below shows how different kinds of bicycles are connected to the generic concept 'bicycle' by means of facets such as 'according to shape', 'according to use', etc.

bicycle
bicycle (according to its shape)
folding bicycle
monocycle
recumbent bicycle
etc.

bicycle (according to its use)
children's bicycle
city bicycle
exercise bicycle
indoor cycling
bike mountain bicycle
etc.
bicycle (according to the way it is propelled)
electric bicycle
hybrid bicycle
standard bicycle

– **Handling different types of relations**

An ideal conceptual structure should account for more than taxonomic relations. To provide a more complete representation of musical instruments, new relations can be added to the taxonomy shown in Figure 8.1. Musical instruments have parts such as 'body', 'strings', and 'keys' and these parts differ from one class of instruments to another. Figure 8.4 shows how specific parts are attached to string instruments. Parts were introduced and organized in the form of a small taxonomy under the generic concept 'part of string instrument'. Then partitive relations are declared between classes of these two taxonomies: more specifically, 'body' is a <part-of> 'string instrument' and 'chin rest' is a <part-of> 'violin'.

The definition of these relations must take into consideration the different levels of the taxonomy and the characteristics assigned to each level. In other words, when 'body' is defined as a <part-of> 'string-instrument', it is inherited by all the specific concepts under 'string instrument', namely 'violin', 'cello', 'double bass', etc. if we had defined 'chin rest' as a <part-of> a 'musical instrument', for instance, it would have been wrongly inherited by all the subordinate concepts of 'musical instrument'.

- musical instrument
  - + woodwind instrument
  - + brass instrument
  - + string instrument
    - – violin
    - – cello
    - – double bass
    - – …
  - + percussion instrument

'body' <is part of> 'string instrument'

'chin rest' <is part of> 'violin'

- part of string instrument
  - – body
    - – upper bout
    - – lower bout
    - – …
  - + neck
  - + chin rest
  - + scroll
  - + string

Figure 8.4 Textual representation of two types of relations for musical instruments

– **Taking characteristics from different sources**

So far, we have assumed that inheritance applies to a situation in which a single generic concept passes on its characteristics to sets of specific concepts. However, there are cases where concepts can lend themselves to different kinds of classifications according to a phenomenon called ***multidimensionality*** (Section 5.1.5). We saw above that facets allow us to introduce different subdivisions in a conceptual structure.

An alternative method used to manage multidimensionality, called ***multiple inheritance***, consists in attaching a specific concept to two or more generic concepts. The specific concept then inherits the characteristics that were assigned to these generic concepts. For instance, we might wish to account for the fact that the concept 'car' <is-a> 'passenger vehicle' and has characteristics defined for this more general concept, i.e. 'used to carry passengers'. We can also account for the fact that 'car' <is-a> 'light vehicle' and inherits the characteristics assigned to this other concept class, i.e. 'has a weight between … and …'

Section 5.1.4 explained that the organization of concepts can be affected by different perspectives. Diverging perspectives on the same concept can be associated with different fields of knowledge or different subdivisions within the same domain. This second form of ***multidimensionality***, can also be handled with multiple inheritance. This is shown in Figure 8.5 which is an illustration of how the concept 'methane' is represented in *EcoLexicon* (2018). 'Methane' is connected to the generic concepts 'greenhouse gas', 'chemical compound', 'hydrocarbon', and 'gas' through a <type-of> relation. Furthermore, in EcoLexicon (2018), users can visualize conceptual structures by limiting their span to a given subdomain (called *contextual domain*) that provides a single perspective on the chosen concept based on how the domain of the environment was subdivided by the designers of the knowledge base.

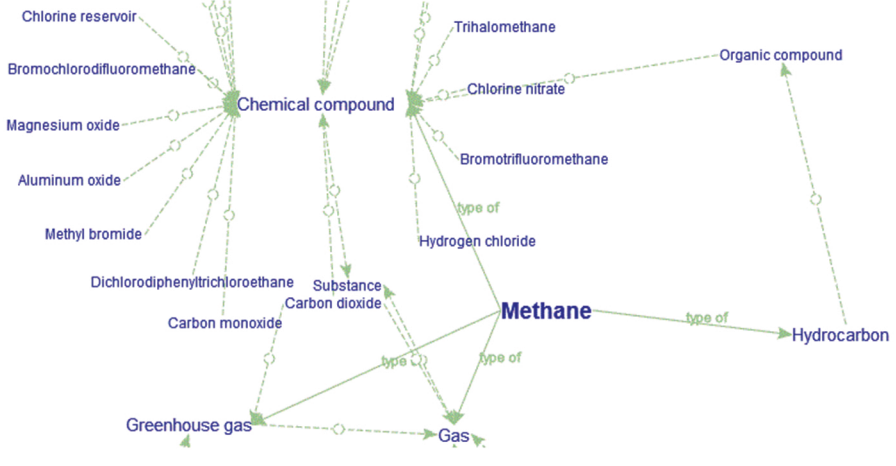


Figure 8.5 'Methane' with multiple <is-a> relations with other concepts (EcoLexicon 2017)

– **Blocking characteristics**

It was assumed so far that all characteristics assigned to a generic concept are necessarily held by specific concepts. However, many cases challenge the full inheritance of characteristics. For instance, let us assume that, in a conceptual structure accounting for bicycles, the following characteristics are used to define the concept 'bicycle': 'small vehicle', 'propelled by human energy', 'has two wheels', 'has two pedals', and 'has a saddle seat'. We might want to consider the concepts 'unicycle' and 'tricycle' as kinds of bicycles and link them to the generic 'bicycle', but the characteristics 'has two wheels' would be wrongly inherited by 'unicycle' and 'tricycle'. Should we create a new class in our hierarchy for 'unicycle' and 'tricycle' or add new subdivisions under the concept 'bicycle' to further organize concepts according to the number of wheels they have?

A more efficient method to tackle this problem consists in ***blocking*** the ***inheritance*** of chosen characteristics. This can be done when most characteristics assigned to a generic concept apply to all but a few subordinate concepts. In our example, we could prevent the 'has two wheels' characteristic to be inherited by 'unicycle' and 'tricycle'.

8.1.3 Linking conceptual structures and definitions

Chapter 5 (Section 5.1.3) explained how terminological ***definitions*** reflect the position of concepts in conceptual structures. Analytical definitions are designed to capture the link between a specific and a generic concept. Furthermore, they specify how specific concepts differ for their neighbors. Definitions are thus less formal versions of conceptual structures. Since there is an obvious relationship between definitions and conceptual systems, authors have suggested methods to link them more directly.

One method suggests using predefined ***templates*** as the one shown in Figure 8.6.[73] The first part of the figure is a definitional template that was designed for a conceptual class called 'hard coastal defense structure'. This template applies to other more specific concepts that belong to this class. The second part of the figure shows the template obtained for the concept 'groin', one of the concepts that belong to the class 'hard coastal defense structure'.

Since it is a specific concept with regard to 'hard coastal defense structure', 'groin' inherits characteristics from the generic concept. Its definition starts by stating that it is a hard coastal defense structure (the genus, indicated in the <type-of> part

73. This proposal was made by León Araúz et al. (2012) for definitions that appear in the terminological knowledge base *EcoLexicon* (2018).

of the template). Then, the different parts of the template defined for the conceptual class – a list of relations labeled <located-at>, <made-of> and <has-function> – are further specified. For instance, the <made-of> relation (that is instantiated by 'material' in the 'hard coastal defense structure' template), is instantiated with more specific forms of material in 'groin', i.e. concrete, wood, metal and/or rock. We can then see how the different parts of the template are connected to the textual definition.

HARD COASTAL DEFENSE STRUCTURE	
type-of	CONSTRUCTION
located-at	SHORELINE
made-of	MATERIAL
has-function	COASTAL DEFENSE

GROIN

Hard coastal defense structure made of concrete, wood, metal and/or rock perpendicular to the shoreline built to protect a shore area, retard littoral drift, reduce longshore transport and prevent beach erosion.

type-of	HARD COASTAL DEFENSE STRUCTURE
located-at	PERPENDICULAR TO SHORELINE
made-of	CONCRETE
	WOOD
	METAL
	ROCK
has-function	SHORE PROTECTION
	LITTORAL DRIFT RETARDATION
	LONGSHORE TRANSPORT REDUCTION
	BEACH EROSION PREVENTION

Figure 8.6 Definitional templates in EcoLexicon (León Araúz et al. 2012, cited in L'Homme and San Martín 2016)

8.2 Structures based on terminological relations

The description of large sets of paradigmatic and syntagmatic relations between terms can help us better understand the ***terminological structure*** of a field of knowledge. The contents of Section 7.2 showed that terminological relations are extremely diversified (synonymy, various forms of antonymy, relations between a predicative unit and arguments or circumstantials, collocations, etc.). They link different types of terms: different parts of speech, terms that denote entities or activities, etc. Relations also differ structurally, i.e. they can connect terms in the

terminological structure – ***paradigmatic relations*** – or link two terms or a term and a lexical unit in a sentence – ***syntagmatic relations***.

Defining relations between terms is quite a different task than classifying concepts in a conceptual structure. While conceptual systems are built in accordance with the way knowledge is organized in a given domain, terminological structures are driven by the meaning of terms. However, it is our contention that terminological structures offer a window, albeit a different one, on knowledge. Furthermore, they constitute a more direct means for observing how this knowledge is expressed in language.

Typically, a structure derived from terminological relations does not offer as orderly a representation as conceptual systems. Moreover, since relations are diversified, different structures can be unveiled depending on: (1) the types of relations taken into consideration; (2) the framework used to represent them. In this section of the chapter, we will focus on two different frameworks and see how they can help us exploit terminological relations and represent them in a meaningful way. The first one is the system of lexical functions proposed by Explanatory Combinatorial Lexicology; the second one is Frame Semantics.

8.2.1 Lexical functions to reveal terminological structures

Terminological structures can be developed on the basis of ***lexical functions*** (LFs) that were briefly introduced in Section 3.4.1 devoted to ***Explanatory Combinatorial Lexicology*** (ECL). Lexical functions were designed for the general lexicon, but we will see that they are especially useful for capturing relations between terms. Section 8.2.1.1 presents some basic notions about lexical functions. Then, Section 8.2.1.2 will explain how they can be applied to semantically related terms to give us valuable information about how these terms are connected. Finally, we will show how they can contribute to obtain a picture of the terminological structure in a given domain.

8.2.1.1 *The workings of lexical functions*

Even if terminological relations are extremely diversified, many are recurrent in the sense that they connect many different terms. For instance, the relation between *deforest* ↔ *deforestation* is identical to the one that appears between *erode* ↔ *erosion*, and *contaminate* ↔ contamination. In all these pairs, terms belong to different parts of speech but convey the same meaning. Similarly, the relation between *emit* and *emitter* is the same as between *program* and *programmer*, as well as between *pollute* and *polluter*. The noun corresponds to the typical **Agent**. Table 8.1 gives additional examples of recurrent relations in the domains of cycling, computing, and the environment.

Standard general language dictionaries already describe some recurrent relations, such as synonymy, near-synonymy and antonymy. More sophisticated resources take into account other paradigmatic relations. For instance, hypernymy and hyponymy are considered to be one of the main structuring relations for nouns in the general language resource WordNet.[74] Figure 8.7 shows how *hammer* (in the sense of 'tool') is connected to a hierarchy of hypernyms. This view places *hammer* within a long chain of hypernyms going all the way to 'entity'.

Table 8.1 Recurrent terminological relations

Term 1	Term 2	Explanation
cycle	*cyclist*	Term 2 designates the typical **Agent** of the activity expressed by term 1
emit	*emitter*	
program	*programmer*	
disinfect	*disinfectant*	Term 2 designates the typical **Means** used to carry out the activity expressed by term 1
fertilize	*fertilizer*	
erode, vi	*erode, vt*	Term 2 adds a meaning of 'cause' to the meaning of term 1
warm, vi	*warm, vt*	
activate	*deactivate*	Term 2 means going back to a prior state (Reversive antonym)
contaminate	decontaminate	
green	*clean*	Term 2 shares most of its meaning with Term 1 (Near-synonym)
human$_A$	*anthropogenic*	
launch	*start*	
file	*create a ~*	Term 2 means 'create' when applied to Term 1
program	*write a ~*	
password	*define a ~*	

Explanatory Combinatorial Lexicology suggests that the lexicon embodies a much broader range of relations, both paradigmatic and syntagmatic.[75] The framework designed a sophisticated system, called ***Lexical functions*** (LFs) (Mel'čuk et al. 1995;

74. The resource also takes into account synonyms and near-synonyms (in what are called *synsets*), meronymy, derivationally related forms and antonymy, depending on the lexical unit described.

75. As was mentioned in Section 3.4.1, ECL accounts for semantic derivation (that corresponds to paradigmatic relations) and collocations (a subtype of syntagmatic relations).

S: (n) **hammer** (a hand tool with a heavy rigid head and a handle; used to deliver an impulsive force by striking)
inherited hypernym
- S: (n) hand tool (a tool used with workers' hands)
  - S: (n) tool (an implement used in the practice of a vocation)
    - S: (n) implement (instrumentation (a piece of equipment or tool) used to effect an end)
      - S: (n) instrumentality, instrumentation (an artifact (or system of artifacts) that is instrumental in accomplishing some end)
        - S: (n) artifact, artefact (a man-made object taken as a whole)
          - S: (n) whole, unit (an assemblage of parts that is regarded as a single entity) *"how big is that part compared to the whole?"; "the team is a unit"*
            - S: (n) object, physical object (a tangible and visible entity; an entity that can cast a shadow) *"it was full of rackets, balls and other objects"*
              - S: (n) physical entity (an entity that has physical existence)
                - S: (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Figure 8.7 *Tool* and its inherited hypernyms in WordNet (2017)

Meľčuk 1996), to capture recurrent relations in different languages. Technically, a lexical function applies to a ***keyword*** and produces a ***value*** or a short list of different values. It is written as follows:

f(x) = y
in which
f stands for a lexical function (LF);
x stands for a keyword; and
y stands for a value produced by the application of the LF to the keyword.

– **A system for capturing many different relations**

Lexical functions are designed to account for many different relations between lexical units. They cover all the relations covered in Section 7.2 and many more. Some common relations are examined below; less common ones are mentioned in following sections.

For example, we can capture the paradigmatic relation between the terms that were mentioned at the beginning of this section with the LF $\mathbf{S_0}$:

$\mathbf{S_0}$(*erode*) = *erosion*
$\mathbf{S_0}$(*deforest*) = *deforestation*
$\mathbf{S_0}$(*program*) = *programming*

S_0(*compatible*) = *compatibility*

S_0 represents the relation between two terms that convey the same meaning, but that belong to different parts of speech. The keyword can be a verb or an adjective; the value is a noun. Similar LFs are associated with other parts of speech. V_0, when the value produced is a verb; A_0, when the value is an adjective; and Adv_0, when the value obtained is an adverb.

Lexical functions can also be used to encode more common relations such as synonymy, near-synonymy, hypernymy, hyponymy and different forms of antonymy that were presented in Section 7.2.3. Two main lexical functions are used to represent ***hypernymy*** and ***hyponymy***, i.e. **Gener** and **Hypo**.[76] We could use them to represent the relations between *hammer* and *tool* depending on which direction the relation is expressed. As was mentioned in Section 7.1.1 these ***relations*** are ***asymmetric***.

Gener(*hammer*) = *tool*
Hypo(*tool*) = *hammer, screwdriver*

Two basic lexical functions distinguish the types of synonyms presented in Section 7.2.1: **Syn** stands for ***exact synonymy*** and **QSyn** stands for ***near-synonymy***. Since these two LFs describe symmetric relations, the same LF should be used regardless of the direction in which the relation is expressed.[77]

Syn(*cardiopathy*) = *heart disease*[78]
QSyn(*habitat*) = *territory*
QSyn(*habitat*) = *environment*
QSyn(*clean*) = *green*
QSyn(*human*) = *anthropogenic*
QSyn(*lauch*) = *start*

---

76. Some hyponyms might be represented with more explicit LFs. For instance, although a *microhabitat* is a kind of *habitat*, it might be more useful to represent it as follows:

AntiMagnTaille(habitat) = microhabitat

In this lexical function, **AntiMagn** stands for 'less' and **Taille** for 'size'. The LF indicates that *microhabitat* is obtained when adding the component 'reduced in size' to *habitat*.

77. ECL also makes use of a specific encoding to indicate that one member of the pair of synonyms has a broader or a narrower meaning ($Syn_{\subset}$ or $Syn_{\supset}$). In some cases, these two lexical functions can be difficult to distinguish from **Gener** and **Hypo**. Another encoding indicates that there is an intersection of meaning ($Syn_{\cap}$).

78. Synonyms can further be labeled according to usage. For instance, *heart disease* could be labeled as common usage. Conversely, *anthropogenic* could be labeled as scientific usage compared to its near-synonym *human*.

LFs can capture most differences between categories of opposites. ***Complementary antonyms***, ***gradable antonyms*** and ***reversives*** are represented with the lexical function **Anti**.[79] ***Conversives*** are represented with the lexical function $\mathbf{Conv}_{ij}$. The subscript numbers indicate which arguments were permuted (the relation between lexical functions and the argument structure of the keyword is explained below). For instance, the relation between *send* and *receive* would be captured as follows:

$\mathbf{Conv}_{31}$(*send*) = *receive*
$\mathbf{Conv}_{31}$(*receive*) = *send*

Finally, ***contrastive terms*** are represented with the lexical function **Contr**. For example, the relation between *ocean* and *continent* would be represented **Contr**(*ocean*) = *continent*. Since the relationship is symmetric, it would be expressed the same way in the other direction (**Contr**(*continent*) = *ocean*).

– **An explicit connection with the argument structure of the keyword**

A large set of lexical functions are designed to link the lexical relation to the argument structure of the keyword. This was already shown above with the $\mathbf{Conv}_{ij}$ LF that was applied to the *send* ↔ *receive* pair of conversive terms. The same relation holds between the verb *run,* defined as **X (a vehicle)** runs on **Y (an energy source)** and *propel*, defined as **X (an energy source)** propels **Y (a vehicle)**. In other words, in order to obtain *propel* from the meaning of *run*, we must change the order of the arguments of *run*. The same procedure must be applied to *propel* to obtain *run*.

$\mathbf{Conv}_{21}$(*run*) = *propel*
$\mathbf{Conv}_{21}$(*propel*) = *run*

Conv states that the arguments of the keyword change positions; the subscript characters "$_{21}$" express how the arguments of the keyword must be ordered to produce the value.

Conv is not the only LF accounting for how lexical relation hinges on the argument structure of the keyword. More examples with terms mentioned in Chapter 7 are listed below.

Name of the first argument of the keyword
$\mathbf{S}_1$(*cycle*) = *cyclist*
$\mathbf{S}_1$(*emit*) = *emitter*
$\mathbf{S}_1$(*program*) = *programmer*

79. Some less direct opposites can also be represented with QAnti (for near-antonyms).

Name of the third argument of the keyword
$\mathbf{S}_3$(*disinfect*) = *disinfectant*
$\mathbf{S}_3$(*fertilize*) = *fertilizer*

In addition to the paradigmatic relations presented so far, LFs can be used to capture the subtle differences between adjectives and verbs as shown below with the terms linked to *erode* (**X** erodes) and pollute (**X** pollutes **Y** with **Z**). These relations were presented in Section 7.2.2.4).

The adjective that applies to the first argument of the keyword
$\mathbf{A}_1$(*erode*) = *eroding*
$\mathbf{A}_1$**Perf**(*erode*) = *eroded*

An adjective that applies to the first argument and conveys the meaning of 'may undergo the process'
$\mathbf{Able}_1$(*erode*) = *erodible*

The adjective that applies to the second argument of the keyword
$\mathbf{A}_2$**Perf**(*pollute*) = *polluted*

The adjective that applies to the third argument of the keyword
$\mathbf{A}_3$(*oxidize*) = *oxidizing*

– **Capturing the properties of syntagmatic relations**

Lexical functions are particularly helpful to account for syntagmatic relations, more specifically the relation between the ***base*** and the ***collocate*** in a collocation. The LF **Magn**, for instance, stands for 'intensification' and produces different collocates depending on the keyword to which it is applied. A lexical function can be associated with more than one value, as shown below.

Magn(*storm*) = *powerful, severe, strong, violent*
Magn(*thunderstorm*) = *intense, severe*
Magn(*agriculture*) = *intensive*
Magn(*biodiversity*) = *high*

Furthermore, as in paradigmatic relations, some recurrent relations can be observed between different sets of bases and collocates that we might want to represent consistently. For example, verbs that express the idea of 'use' in the field of computing differ according to the term selected.

'use' Internet: surf the ~
'use' page: visit a ~
'use' file: edit a ~
'use' mouse: move a ~
'use' password: use a ~

However, collocations require a more complex apparatus than most paradigmatic relations, since, lexical units or terms also share a syntactic relation within a collocation. Hence, many LFs designed to represent syntagmatic relations also encode information about their syntactic structure. In fact, they encode three different properties that are explained below.

A. The syntactic relation between the keyword and the collocate. For example, verbal collocates of *mouse*, namely *move* and *click*, are distinguished according to the position of *mouse*. When used in combination with *move*, *mouse* is a direct object; when used in combination with *click*, *mouse* is a complement

move a mouse: verb + mouse
click on … (e.g. an icon) with a mouse: verb … + mouse

B. The argument structure of the keyword. Many collocations, especially verbal and nominal collocations, convey meanings that involve arguments of the keyword. As can be seen in Figure 8.8 with the collocates of *mouse*, *move a mouse* involves the first argument of *mouse*, which corresponds to the person who uses it. The meaning of *click on something with the mouse*, involves both arguments. The first argument of *mouse* carries out the clicking; the second argument of *mouse* instantiates the object on which the clicking is done. Hence, the lexical function that encodes *move*, with respect to *mouse*, must point to the first argument; the one that encodes *click* must point to both arguments. Specific LFs used for these two examples are given further on.

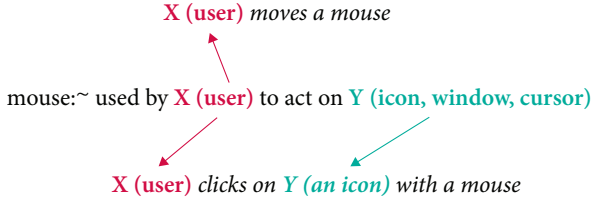


Figure 8.8 Argument structure of *mouse* and two collocations

Another example can be given to illustrate the link between collocations and the argument structures of keywords. Two different verbal collocates can be used with the term *risk* (risk caused by **X** to **Y** that results in **Z**: *RISK of adverse impact from climate change; RISK in region*): *pose* (*pose a risk*) and *be* (*be at risk*). *Pose* and *be* are *support verbs* that are empty semantically. However, their

choice depends on the argument that controls them: *pose* applies to the first argument of *risk; be* applies to the second as shown below.

> **X** *poses a risk* (*frequency of severe weather events pose a* RISK *to several areas*)
> **Y** *is at risk* (*semi-arid regions are at greater* RISK)

C. The general and abstract meaning of the collocate. In a collocation, the meaning of collocates are considered with respect to the keyword they are combined with. For example, both *move* and *click* refer to typical uses of the mouse. Other collocates mentioned earlier convey a meaning of 'creation', e.g. *create a file, a password, develop a program*. Other collocates such as *pose* and *be* with *risk* are empty semantically.
To encode the respective meanings of *move* and *click* when used with *mouse*, the following LFs are used.

> **Real**$_1$(*mouse*) = *move a* ~
> **Labreal**$_{12}$(*mouse*) = *click on … with a* ~

Real$_i$ and **Labreal**$_{ij}$ are used to represent the meaning associated with typical uses or the inherent activity associated with the keyword. **Real**$_i$ serves to encode collocations in which the keyword is first complement; **Labreal**$_{ij}$ accounts for collocations in which the keyword is second complement. The subscript numbers stand for the arguments involved in the collocation. Compare the following encodings for different collocations. A third LF is used to represent 'inherent realization of' for keywords that appear as subjects, i.e. **Fact**$_i$.

> **Real**$_1$(*Internet*) = *browse the* ~
> **Labreal**$_{12}$(*Internet*) = *search for … on the* ~
> **Labreal**$_{12}$(*keyboard*) = *enter … on a* ~
> **Real**$_1$(*habitat*) = *inhabit a* ~

> **Fact**$_0$(*program*) = *the* ~ *runs*
> **Fact**$_2$(*mother*) = *the* ~ *raises …*
> **Fact**$_2$(*biologist*) = *the* ~ *studies …*

– **How many lexical functions?**

Theoretically, there could be as many lexical functions as there are relations in the lexicon. In practice, however, LFs – at least standard LFs – are suggested for recurrent relations in languages.

Approximately 60 lexical functions are designed to encode simple paradigmatic and syntagmatic relations. Furthermore, LFs can be combined to express more complex meanings. An example was mentioned earlier with **AntiMagn: Magn** expresses intensification and **Anti** represents opposition. Other examples can be given with the LFs **Fact**$_i$, **Real**$_i$ and **Labreal**$_{ij}$. They can be combined with

LFs for processual meanings, i.e. **Incep** 'begin', **Cont** 'continue' and **Fin** 'end'. Examples of simple and complex LFs applied to English and French terms are given in Table 8.2.

Lexical functions can be used to encode recurrent and general meanings and/or relations, such as the ones presented in Table 8.2. However, some meanings are much less common, and we might still want to represent them in resources. Furthermore, the system of standard lexical functions – about 60 LFs and combinations of simple LFs into complex ones – was created on the basis of the general lexicon. Other kinds of phenomena might be recurrent in specialized domains. In these two situations, non-standard lexical functions can be created or a non-standard component can be added to a standard LF.

Table 8.2 Examples of specialized collocations encoded with lexical functions

Term acting as keyword	Lang.	Collocation	Lexical function	Short description of the LF for the collocation used as example
algorithmique	*Fr*	*faire de l'~*	**Oper$_1$**	Support verb when keyword is 1st complement
base de données	*Fr*	*construire une ~*	**Caus$_1$Func$_0$**	Cause to exist; create
blogue	*Fr*	*publier … dans un ~*	**Caus$_1$Func$_3$**	Cause something to be in relation with the 3rd argument
character string	*En*	*delete a ~*	**Liqu$_1$Func$_0$**	Cause to cease to exist; eliminate
courriel	*Fr*	*écrire un ~, taper un ~*	**Caus$_1$Func$_0$**	Cause to exist; create
file	*En*	*create a ~*	**Caus$_1$Func$_0$**	Cause to exist; create
glacier	*En*	*~ retreats*	**IncepPred Minus**	Start being less important; diminish
habitat	*En*	*inhabit a ~*	**Real$_1$**	Use
Internet	*En*	*connect to the ~, go on the ~*	**IncepReal$_1$**	Start using
Internet	*En*	*browse the ~*	**Real$_1$**	Use
Internet	*En*	*disconnect from the ~*	**FinReal$_1$**	Stop using
population	*En*	*~ declines*	**IncepPred Minus**	Start being less important; diminish
programmeur	*Fr*	*bon ~*	**Ver**	Is as should be
species	*En*	*~ survives*	**ContFunc$_0$**	Continue existing
species	*En*	*~ becomes extinct*	**FinFunc$_0$**	Start to no longer exist; disappear, die
storm	*En*	*severe ~, strong ~, violent ~*	**Magn**	Intense

For instance, in the field of computing, many verbal and nominal collocates express a form of transformation as shown in the examples below. These collocations would require a special kind of encoding such as a non-standard LF.

alias: change an ~
application: update an ~
character: encode ~s
character string: convert a ~
database: export a ~
line; edit a ~
message: encrypt a ~
table: import a ~

8.2.1.2 *Lexical functions for terminology*

It should be obvious by now that lexical functions are useful to encode relations, both paradigmatic and syntagmatic, in terminological resources. The system of lexical functions can be further exploited in many different ways. This section examines more closely some terminological phenomena that LFs contribute to reveal in specialized domains. Other more sophisticated uses and extensions of lexical functions are presented in further sections.

– **Explaining multiple relations shared by a term with other terms in a specialized domain**

Lexical functions label the many relations that one term shares with others in a specialized field of knowledge and thus distinguish relations in a formal and consistent way. Figure 8.9 is a graphical representation of the relations held by *pollute*$_{1b}$ (**X** pollutes **Y** with **Z**) with other terms in the environment.[80]

LFs provide a formal distinction for the relations shared by the nouns *polluter* and *pollutant* with *pollute. Polluter* is the typical name of the first argument of the verb (*POLLUTERS should pay for the pollution they cause*) and is encoded with $\mathbf{S_1}$. *Pollutant*, on the other hand, is the name of the third argument (*ozone is a common air POLLUTANT*) and is encoded with $\mathbf{S_3}$. Similarly, LFs differentiate the two adjectives *polluted* and *polluting*. The first one – represented with $\mathbf{A_2}$**Perf** – applies to the second argument (*POLLUTED water also contains viruses*); the second one – represented with $\mathbf{A_3}$ – to the third argument (*emissions of other POLLUTING gases and particles*).

---

80. This figure shows that two meanings for pollute were identified. *Pollute*$_{1a}$ has two arguments (**X** pollutes **Y**: *Hazardous wastes may POLLUTE the soil, air, surface water, or ground water)*; whereas *pollute*$_{1b}$ has three arguments (**X** pollutes **Y** with **Z**: *we are POLLUTING the atmosphere with toxic emissions*). The verbs are semantically related and should be linked for this reason.

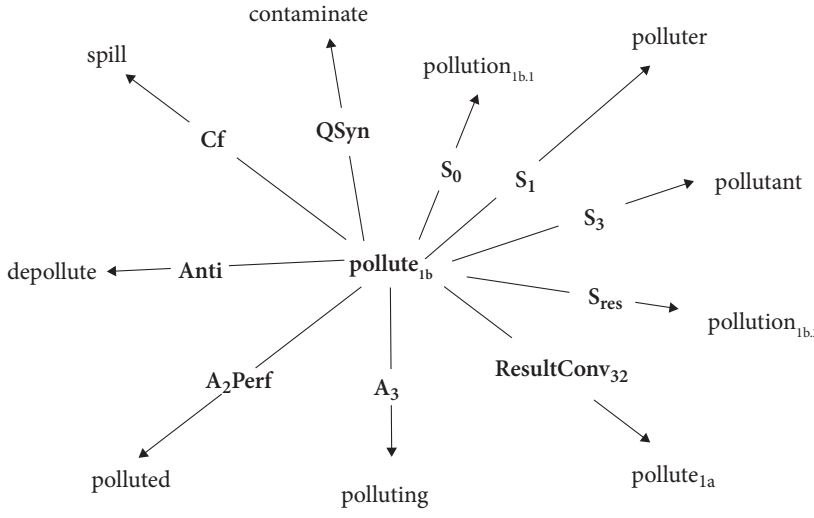


Figure 8.9 Relations shared by *pollute*$_{1b}$ with other terms encoded with lexical functions

– **Making semantic distinctions more explicit**

When lexical items carry multiple meanings, each one is connected to different sets of units. Lexical functions allow us to readily perceive semantic distinctions by attaching different sets of terms to polysemous items and by providing an explanation for the relation. This is shown in Figure 8.9 where the two meanings of *pollution* (*pollution*$_{1b.1}$ 'activity' and *pollution*$_{1b.2}$ 'result') are linked to pollute through distinct LFs (S_0 and S_{res}). Another example is given below with the noun *litter*. It conveys three different meanings that are relevant for the environment. Each meaning leads to the establishment of different relations.

> *litter*$_1$, *n.*: *a* ~ of **Y** produced by **X** ... *a* LITTER *of one to five young (usually two) are born in January or February of the following year.*
> *(Mississippi Museum of Natural Science 2014)*
>
> **Name for Y (S_2)** *young, pup, kitten, cub, offspring*
>
> **Name for X (S_1)** *animal, female*
>
> **Inherent realization carried out by the first argument ($Real_1$)**
> *give birth to a ~, have a ~*
>
> *litter*$_2$, *n.*: ~ containing **X** in **Y** *The leaf* LITTER *can provide adequate levels of moisture and food resources for amphibians to carry out foraging.*
> *(Meza-Parral & Pineda 2015)*

Name for X (S_1) *leaf*

Name for Y (S_2) *forest, soil*

Inherent realization applied to the second argument ($Fact_2$)

~ covers …

litter$_3$, n.: ~ containing **X** in **Y** *Most of the* LITTER *that ends up in our ocean is lightweight, durable, strong, inexpensive and long-lasting plastic. Standardize presentation of examples* (*Hardesty et al. 2015*).)

Name for X (S_1) *plastic, debris*

Name for Y (S_2) *environment, beach, ocean, water*

Related meaning (Cf) *waste*

Adjective expressing a typical location ($A_{@}Sloc$)

marine ~

Noun expressing the activity that consists in eliminating ($S_0Liqu_{@}Func_2$)

removal of ~ from …

– **Encoding meanings and/or relations consistently**

Lexical functions label the same meaning or the same relation consistently, even when they involve different pairs of terms. This feature allows us to get a global view on relations shared by large sets of terms in specialized domains. For instance, we could retrieve from a terminological resource that encodes terminological relations with LFs, all the pairs of terms that share a relation similar to that shared by *pollute$_{1b}$* and *pollution$_{1b.2}$*. Figure 8.9 shows that this relation is encoded with the LF **Sres**. Many more pairs of terms share this relation in the environment: *emit* ↔ *emission; fragmentation* ↔ *fragment; to harvest$_V$* ↔ *harvest$_N$; to plant* ↔ *plantation*, and so on.

Other examples were given in Section 8.2.1.1 with the meaning of 'use' that can be expressed by different verbs depending on the terms to which it applies. An even more spectacular example can be given with the meaning of 'create' in the field of computing (represented with the LF $Caus_iFunc_0$). This meaning can be expressed with many different verbs as shown below.

$Caus_1Func_0$(*alias*) = *create, define*
$Caus_1Func_0$(*code*) = *write*
$Caus_1Func_0$(*condition*) = *specify*
$Caus_1Func_0$(*file*) = *create*
$Caus_1Func_0$(*folder*) = *create*
$Caus_1Func_0$(*interface*) = *develop*
$Caus_1Func_0$(*page*) = *build*
$Caus_1Func_0$(*parameter*) = *define*
$Caus_1Func_0$(*password*) = *create, define, set*
$Caus_1Func_0$(*program*) = *write*
$Caus_1Func_0$(*site*) = *construct, create*

– **Linking relations to argument structures and distinguishing them more precisely**

As was mentioned earlier, many lexical functions establish a connection between the argument structure of the keyword and paradigmatic and syntagmatic relations. This becomes particularly important when, for the same keyword, related terms express close meanings but apply to different arguments. This was shown in Figure 8.9: the relations between *pollute*$_{1b}$ with *polluter* and *pollutant* can be explained depending on the argument they stand for (first argument for *polluter* and third argument for *pollutant*). Similarly, the two adjectives *polluted* and *polluting* apply to the second and third arguments of *pollute*$_{1b}$ respectively.

The same principle allows us to distinguish syntagmatic relations. Consider an example taken from a business dictionary that we mentioned in previous chapters, i.e. Binon et al. (2000). In this dictionary, the connection between syntagmatic relations and the argument structure is established via a system of ***variables*** similar to those used in Explanatory Combinatorial Lexicology (see Section 6.5). The variables appear in definitions as well as in the short explanations given for verbal collocates. This is illustrated below with the French term *bien* (*asset*, *goods*). Some verbal collocates express activities carried out by the producer of the goods (**X**) while the others are associated with activities carried out by the user of the goods (**Y**) (Binon et al. 2000: 62–63).

> BIEN: 1.1 Objet matériel ou immatériel réalisé par un agent économique (une entreprise – **X**) et qui sert à la production (par ex. une machine) ou qui est destiné à la satisfaction des besoins des consommateurs et des entreprises (**Y**) (par ex. un téléviseur).
>
> **X** produire un bien, vendre un bien, offrir un bien
>
> **Y** demander un bien, acheter un bien, acquérir un bien, consommer un bien[81]

A second example is given below with the term *password* in the field of computing. It is assumed that the argument structure of the quasi-predicative term is defined as follows:

---

81. GOODS: 1.1 Material or non-material object created by an economic agent (a company – X) and that is used for production (e.g. a machine) or to fill the needs of consumers and companies (Y) (e.g. a television set).

> **X** produce goods, sell goods, offer goods
>
> **Y** request goods, buy goods, acquire goods, use goods

A password: ~ given to **X** by **Y** to act on **Z**

X (1st argument) stands for the entity that needs to access something
Y (2nd argument) stands for the entity responsible for granting access to something
Z (3rd argument) stands for the entity that can be accessed

Different verbs are used with *password* to express its creation or its use. Some collocates are associated with the first argument (**X**), and others with the second (**Y**) or third (**Z**) arguments. Finally, some verbs can be associated with two different arguments. The table below shows how this is made explicit with lexical functions.

A password: ~ given to **X** by **Y** to act on **Z**

enter, type a ~	1st argument	**IncepReal**$_1$ (start using)
ask for a ~, prompt for a ~	2rd argument	**IncepReal**$_2$ (start using)
use a ~	1st argument	**Real**$_1$ (use)
access … with a ~	1st and 3rd arguments	**Labreal**$_{13}$ (use to act on something)
protect … with a ~	2nd and 3rd arguments	**Labreal**$_{23}$ (use to act on something)

– **Revealing the terminological structure of a given domain**

Figure 8.4 provides a first view on the terms that are directly connected to *pollute*$_{1b}$. However, each of the connected terms can also share relations with other environment terms, as shown in Figure 8.10 with *polluting*.

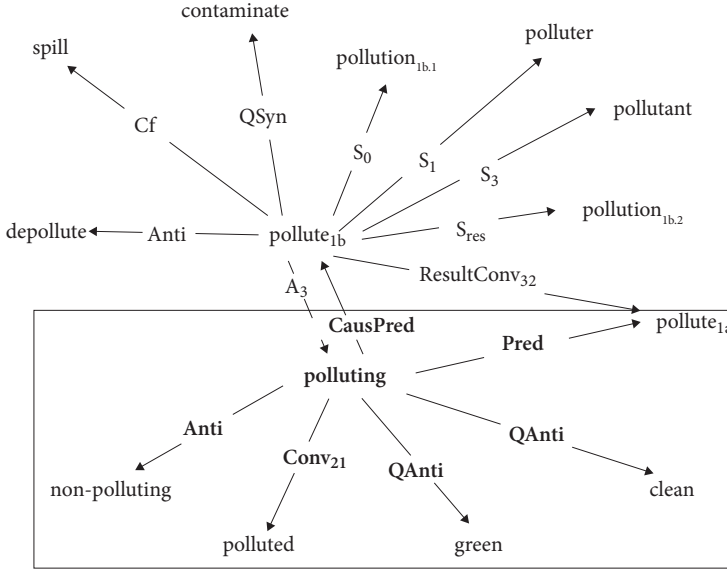


Figure 8.10 Relations shared by *pollute*$_{1b}$ and *polluting* with other terms encoded with lexical functions with a closer view on *polluting*

We could also use one of the terms linked to *polluting* (*polluted*, *green* or *clean*) as a starting point to discover new sets of related terms to gradually discover the structure that links pollutants, contaminants, different ways they affect the environment, and potentially ways to eliminate them.

Similarly, if we look at the relations between the term *animal* with other terms in the field of endangered species, we can further our understanding of the realities surrounding animals in that particular domain. For example, categories of animals can be established according to different features (status, morphology, reproduction, etc.). We can also associate specific stages of life to animals, find out that this life might be threatened, and so on (Table 8.3).

Table 8.3 Terms related to *animal*

Explanation of the relation	Lexical function	Related term(s)
Contrastives	**Contr**	*fungus, plant*
Types of a.	**Hypo**	*bird, fish, insect, mammal, reptile*
That is large in size	**MagnTaille**	*large ~*
That can stop existing	**Able$_1$FinFunc$_0$**	*endangered ~, imperiled ~, threatened ~*
The a. exists	**Func$_0$**	*the ~ lives*
Noun for the a. continues to exist	**ContFunc$_0$**	*survival 1 of the ~*
The a. exists in a given location	**Func$_@$[@:lieu]**	*the ~ inhabits …, the ~ lives in …, the ~ occupies …*
Noun for the a. becomes less important	**S$_0$IncepPredMinus**	*decline of the ~,*
Noun for the a. starts to no longer exist	**S$_0$IncepFinFunc$_0$**	*~ extinction*
Someone or something keeps a a. in its current state	**Caus$_@$ContPredVer**	*preserve an ~, protect an ~*
Set of a.	**Mult**	*community, fauna, population*

Building on some of the terms that are linked to *animal*, we can discover additional relevant relations with respect to endangered species. First, looking at one of the hyponyms of *animal*, namely *bird*, we find out that *bird* itself has hyponyms: *eagle, owl, pelican, peregrine falcon, plover.* The term *bird* is also associated with adjectives denoting vulnerability (*endangered ~, imperiled ~, threatened ~*). We further discover, using *endangered* as a starting point, that other terms appear with this collocate: *endangered bird, endangered fish, endangered mammal, endangered organism, endangered plant, endangered species, endangered subspecies, endangered*

taxon. We could also find out that sets of terms are associated with nouns and verbs denoting different activities or stages of life and existence (*survival of the species, population inhabits* …). Finally, using *inhabit* or *occupy* as starting points, we obtain a new list of terms that denote places where animals live (*inhabit a site, inhabit an area, inhabit an environment, inhabit a reserve*, etc.) (Figure 8.11).

As in the example with *pollute*, we can gradually reveal the structure that connects animals with other entities that interact with them (e.g. plants, organisms, humans), as well as activities that they carry out or situations in which they find themselves. Of course, many steps are required to retrieve this information on the basis of binary relations. In the examples given in the previous paragraph, two to three steps are necessary. In addition, if all terms and the relations they share are considered, the network of relations will soon become quite complex and difficult to manage. Nevertheless, when relations are established, they can be exploited to get a more complete perspective on a domain.

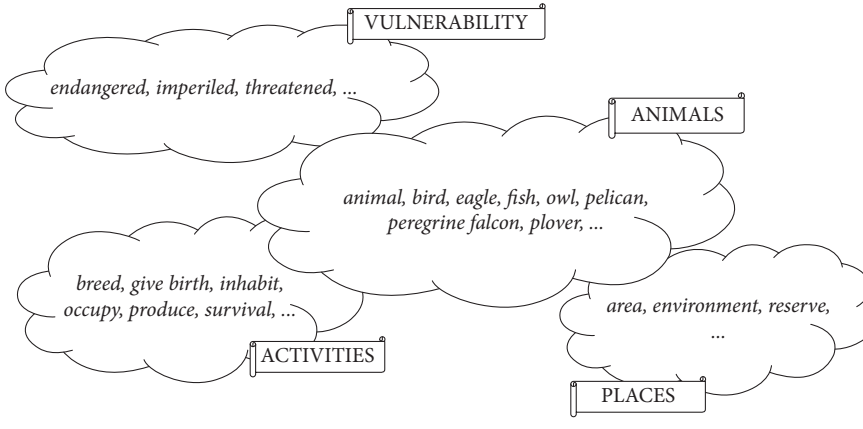


Figure 8.11 Relations between *animal* and other terms linked to endangered species

We will see further on (Section 8.2.2) that Frame Semantics can be used to discover a different kind of terminological structure. It can also be used as an alternative to the representation of binary relations. First, however, we will examine other exploitations of lexical functions that can help us take full advantage of terminological relations.

8.2.1.3 *Exploring terminological relations with "softer" versions of lexical functions*
From the point of view of encoding relations in a terminological resource, lexical functions have several advantages, many of which have been mentioned in previous pages. They take into account three different properties (syntactic,

semantic and argument structure) and provide for a rigorous classification of related terms.

Despite their ability to capture properties of relations, or maybe because of this, LFs have been criticized with regard to their lack of transparency. Indeed, they can be quite difficult to decipher for users who are not familiar with them (which in fact represents nearly everyone but a small group of *aficionados*). For this reason, alternative systems were developed to help users take advantage of their encoding possibilities without having to go through a long learning process. Mel'čuk and Polguère (2007) designed a series of paraphrases in French that can be superimposed on LFs to make them accessible to users.

Two terminological resources, the DiCoEnviro (2018) and the DiCoInfo (2018), use a system of natural language explanations very similar to those proposed by Mel'čuk and Polguère (2007).[82] The original version was adapted in order to account for relations in specialized domains.

Although formulated in natural language, explanations highlight properties of relations in a way that is reminiscent of lexical functions. First, the connection between the related term and the argument structure of the keyword is made explicit by repeating the labels used for arguments in the explanation. This is shown in Figure 8.12 with Web_1 and the collocations *browse the ~, access the ~, search … on the ~*. These collocations are encoded as $\mathbf{Real}_1$, $\mathbf{IncepReal}_1$, and $\mathbf{LabReal}_{12}$ respectively.

Web is defined as a quasi-predicative term with two arguments which were labeled with the typical terms *user*, *information and site*. The collocations *browse the ~* and *access the ~* involve the first argument of *Web*. This is shown explicitly in the explanations accompanying the collocations since the typical term *user* is highlighted with the color associated with it in the argument structure, i.e. red. The collocation *search … on the ~* involves both arguments. The *user* is the argument doing the searching and *information* and *site* represent the objects that are searched (two typical terms were chosen for the second argument since no hypernym includes both). Again, this is made explicit with the repetition of the typical terms and colors associated with them. The syntactic structures of the collocations are stated in the value where the respective positions of the keyword and collocates are reproduced. Finally, the general meaning of the collocation is captured in the

82. The paraphrases developed for the DiCoEnviro and DiCoInfo are superimposed on lexical functions. Hence, in the resources, terminological relations are encoded by terminologists once with LFs and once with natural language explanations. The explanations are displayed in the online textual version of the resources as they can be visualized by users. LFs are used in other instances: we will see further on that they can serve to retrieve translations of collocations.

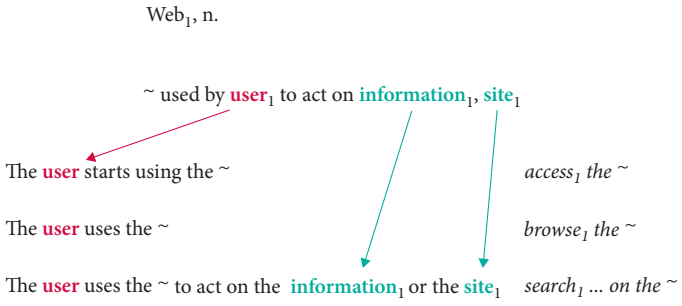


Figure 8.12 Argument structure and explanations for relations

explanation. Since all these relations express uses associated with *Web*, *use* appears in all three explanations (*starts using*, *uses*).

In contrast to LFs that are language-independent, explanations must be adapted if relations are encoded in different languages. This is shown in Table 8.4 for English, French and Spanish.

Table 8.4 Related terms, lexical functions and paraphrases

Related term	Lexical function	Paraphrase
species$_1$		
the ~ survives$_1$	**ContFunc**$_0$	The s. continues to exist
the ~ disappears$_1$	**FinFunc**$_0$	The s. ceases to exist
espèce$_1$		
l' ~ survit$_1$	**ContFunc**$_0$	L'e. continue d'exister
l' ~ s'éteint$_1$	**FinFunc**$_0$	L'e. cesse d'exister
keyboard$_1$		
enter$_1$ … on a ~	**Labreal**$_{12}$	The user uses a k. to act on the data
clavier$_1$		
entrer$_1$ … au ~	**Labreal**$_{12}$	L'utilisateur utilise un c. pour intervenir sur les données
teclado$_1$		
introducir$_1$ … a través de un ~	**Labreal**$_{12}$	El usuario utiliza un t. para operar con los datos

8.2.1.4 *Further classifying relations*

There are about 60 simple lexical functions that encode both paradigmatic and syntagmatic relations. Furthermore, LFs can be combined and non-standard LFs can be created for less recurrent relations. In resources that account for all

possible relations between terms, this variety can soon become overwhelming. For instance, in the DiCoInfo terminological resource, the noun *file* is associated with over 100 related terms. Again, different proposals were made to further classify relations and the LFs used to encode them. A strategy that was implemented in the DiCoEnviro and the DiCoInfo is presented in this section.

First, as was seen in Chapter 7, relations can be classified in broader categories, such as paradigmatic versus syntagmatic, opposites versus synonymy and near-synonymy, and so on. In the DiCoEnviro and DICoInfo, broad categories called ***families*** were defined in order to group sets of relations that share similar properties. Up to five different families are listed in entries.

– Related meanings: This family contains near-synonyms,[83] hypernyms and other terms that belong to the same paradigm, such as co-hyponyms.

mammal: n
Hypernym: animal
Related meanings: bird, fish

copy: vt
Related meanings: paste, copy and paste

– Opposites: This family accounts for all types of opposites, i.e. antonyms, conversives and contrastives, presented in Section 7.2.2.3.

soluble: adj
Antonym: insoluble

install: vt
Reversive: uninstall

– Word family: This family contains words that are semantically related but that often belong to different parts of speech.

download$_{1}$*: vt*
Noun: download$_{1.1}$
Adjective 'that can be + verb': downloadable

erode$_{1a}$*, vi*
Verb 'cause to': erode$_{1b}$
Noun: erosion
Adjectives: eroded, erodible, eroding

83. Related meanings should also include exact synonyms. However, in the resources, exact synonyms are placed in a separate data category. This allows us to highlight their special status with respect to the headword as we will see further on.

– Types of: In this family, terminologists encode all syntagmatic and paradigmatic relations that express a more specific meaning with respect to the headword. Hence, hyponyms and collocations can be included.

environment: n
That concerns a specific location: aquatic ~, coastal ~, marine ~

file, n.
According to the data it contains: data ~, text ~, attachment
According to what it is used for: configuration ~, installation ~

– Combinations: This family contains all verbs and nouns that combine with the term to form collocations.

file, n.
Create or eliminate: create a ~, ~ creation, delete a ~, generate a ~
Use: load a ~, open a ~, edit a ~

habitat, n.
Use: occupy a ~, inhabit a ~, introduce … in a ~
Change or destroy: alter a ~, degrade a ~, degradation of a ~, destroy a ~

– Others: All related terms that cannot be encoded in the previous families are listed in this one. This includes meronyms, holonyms, circumstantials and typical arguments, typical locations, among others.

computer, n.
Parts: motherboard, hard disk, processor, memory

ocean, n
Part: surface
Domain: oceanography

This distribution of lexical relations in different families can be exploited in the presentation of relations to users. In the online versions of the DicoEnviro and the DiCoInfo, families are presented in a table in which related terms are classified according to the family to which they were assigned. Explanations also accompany each related term (see Section 8.2.1.3).

An alternative display of relations is accessible in another resource, called *The NeoVisual* (2018, L'Homme et al. 2018), which distributes related terms around a central one according to the family in which they were classified. Figure 8.13 shows how related terms are clustered around *carbon dioxide*. In the graph, features make some distinctions more readily available than in the table mentioned above. Clusters are differentiated with colors:

– Red: exact synonyms.[84]
– Blue: other related meanings, i.e. near-synonyms, hypernyms, etc.
– Purple: word families.
– Light blue: types of.
– Green: verbal and nominal collocations.
– Light green: other types of relations.
– Burgundy: equivalents in other languages.

Additional colors are also used for other kinds of distinctions:

– Pink: opposites.
– Orange: arguments (see Figure 8.15 with the term *pollute*$_{1b}$).

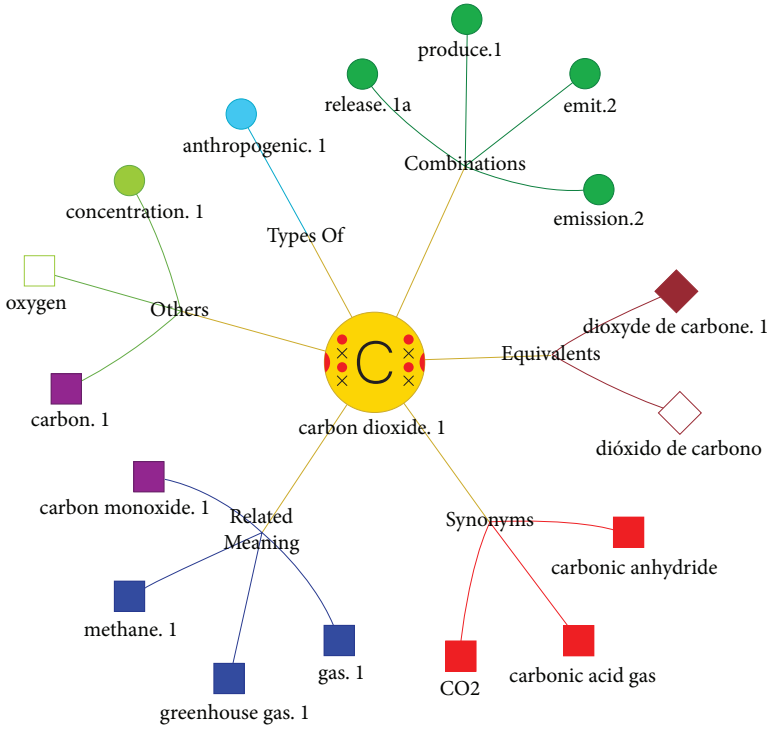


Figure 8.13 Relations held by *carbon dioxide* with other terms (NeoVisual 2018)

84. Since, all other related features represented in the graph are valid for exact synonyms (see Section 7.2.2.2), when clicking on a node where an exact synonym is mentioned, the red lines are highlighted in the graph indicating that links are propagated to synonyms.

In addition to colors, the NeoVisual implements other graphical devices to make properties of relations explicit:

- Square-shaped nodes: paradigmatic relations (although most families contain either syntagmatic or paradigmatic relations, some contain both types and nodes allow users to visualize the distinction at once).
- Triangles: arguments.
- Circles: syntagmatic relations.
- Diamond-shaped nodes: equivalents in other languages.

More precise information can be obtained for a specific relation when selecting an edge. The explanation followed by the lexical function appears in a pop-up. This is shown in Figure 8.14 with the terms linked to *erode*$_{1b}$.

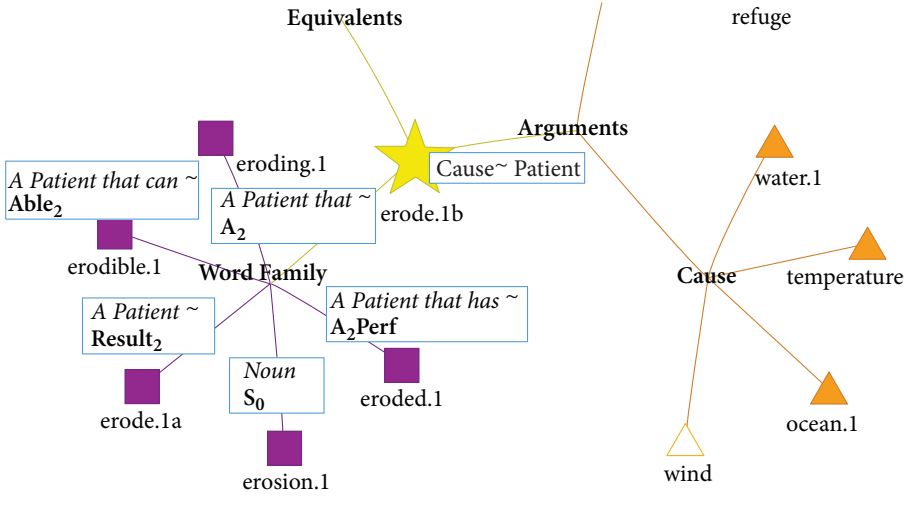


Figure 8.14 Explanations and lexical functions for some related terms given for *erode*$_{1b}$ (NeoVisual 2018)

The NeoVisual is designed to give a first general overview of the relations shared by a chosen term with others. However, users might wish to further explore the terminological structure of the domain. Additional navigational features can be used to gradually reveal other parts of this structure. Each node in a graph can be selected to generate a new graph. For instance, selecting an equivalent will result in displaying the relations shared by this equivalent with other terms in a different language. Users can go back to previously generated graphs by selected a node in the history since the tool keeps track of the different terms that users selected during a session.

Finally, an interesting graphical feature can be used to account for the fact that some related terms can appear in two different families. For instance, *depollute*$_1$ is an antonym with respect to *pollute*$_{1b}$ but it is also morphologically related. Hence, the term *depollute*$_1$ should be placed in the "Opposites" family. The NeoVisual accounts for this "double" relationship as can be seen in Figure 8.15. Instead of adding new edges, the NeoVisual colors the nodes in purple so that users can visualize both types of information, i.e. the family in which the related term belongs and the morphological similarity. In some cases, terms that are morphologically related appear in many different families, including Word family. This is the case with terms linked to *pollute*$_{1b}$ as can be seen in Figure 8.15.

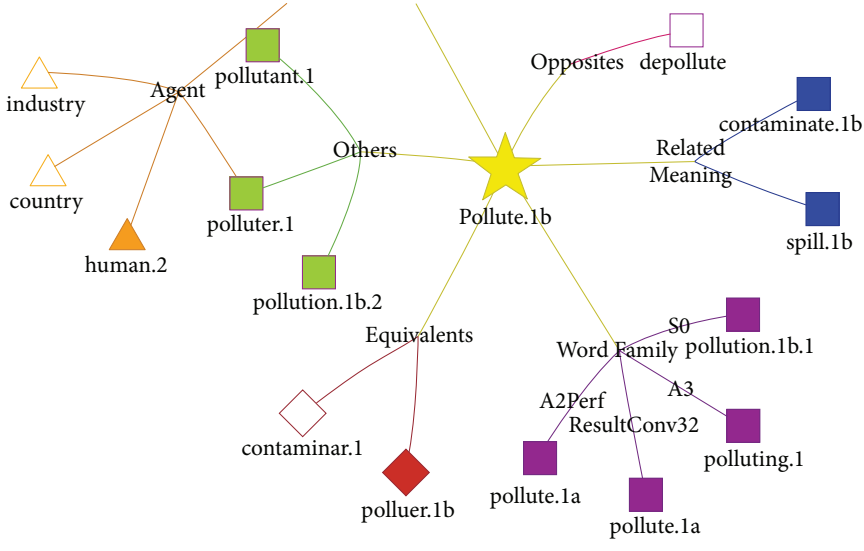


Figure 8.15 Highlighting related terms with morphological similarity (NeoVisual 2018)

8.2.1.5 *Definitions based on terminological structures*[85]

Section 8.1.4 explained how conceptual structures could be exploited to produce more consistent definitions. This section shows how terminological structures can also serve to predict and standardize the contents of definitions.

85. For this section, I am grateful to Antonio San Martín with whom I had and continue to have discussions on the topic of definitions. Much of what is stated in this section is a result of work we carried out together.

For instance, let us assume that the argument structure and the definition of the verb *click*$_1$ runs as follows:

> **click**$_1$, vt: user ~ on icon with mouse
> Definition: A user selects an icon by pressing and releasing a button of a mouse.

The noun *click* and the adjective *clickable* are closely linked to the verb *click*. The adjective can be defined as 'that can be + the meaning of click'. This relation is encoded with the LF **Able**$_2$. The difference between the verb *click* and the adjective *clickable* can be explained by adding the component 'that can be' to the definition of the verb, as shown below.

> **clickable**$_1$: ~icon$_1$
> Definition: An icon that can be selected by a user by pressing and releasing a button of a mouse.

The pattern that was described with *click*$_1$ and *clickable*$_1$ can be applied to all other adjectives that share an **Able**$_2$ relation with a verb. This is shown below with *download* ↔ *downloable*, but the pattern can be used for many other pairs or terms. Other rules can be devised for other regular relations, as shown in Table 8.5.

> **download**$_{1:}$ user ~ application, file from computer, network to computer
> Definition: A user transfers a file or an application from a remote computer or a network and places the file or the application on the computer of the user.

> **downloadable**$_{1:}$ ~ application, file
> Definition: An application or a file that can be transferred by a user from a remote computer or a network and placed on the computer of the user.

In some cases, the application of the rule is straightforward (as with the relations listed in Table 8.5). In others, however, parts of the definitions must be handled with specific constraints or edited. For instance, even if many nouns express the same meaning as verbs (a relation encoded with the lexical function S_0), the choice of the genus in the noun definition depends on the type of event it denotes (a state, an action, a fact, an activity, a process or an event[86]).

> **click**$_{1:}$ A user selects an icon by pressing and releasing a button of a mouse.
> **click**$_{1.1:}$ Action in which a user selects an icon by pressing and releasing a button of a mouse.

86. This classification of meanings is based on an adaptation by Van Valin (2005) of Vendler's verb classification (1967).

In this case, the genus *action* was chosen since the meaning expressed involves a change of state, an agent, and does not have an internal duration.

crash$_{1a}$: A computer or a program stops functioning and responding.
crash$_{1a.1}$: Event during which a computer or a program stops functioning and responding.
In this case, the genus *event* was chosen since the meaning expressed involves a change of state, does not involve an agent and does not have an internal duration.
pollute$_{1b}$: Humans or activities cause a substance to damage an area by entering this area and accumulating in this area.
pollution$_{1b.1}$: Activity in which humans or activities cause a substance to damage an area by entering this area and accumulating in this area.
In this case, the genus *activity* was chosen since the meaning expressed involves a change of state, an agent, and has an internal duration.

Table 8.5 Examples of definitions produced on the basis of regular relations

Term 1	Lexical function	Term 2	Instruction
write$_1$	**De_nouveau**	Rewrite	Add *once again* at the end of term 1 definition.
write$_1$: A drive, a, program or a processor records data in a memory or on a storage device. rewrite$_1$: A drive, a, program or a processor records data in a memory or on a storage device once again.			
melt$_{1a}$	**Caus@**	melt$_{1b}$	Change term 1 definition into a subordinate clause of the main clause whose subject is the instantiation of the first argument of term 2 and whose verb is *cause.*
melt$_{1a}$: Ice changes from solid to liquid state. melt$_{1b}$: Temperature causes ice to change from solid to liquid state.			
migrate$_1$	**A$_1$**	migrating$_1$	Change term 1 definition into a relative clause whose head is the instantiation of the first argument of term 1.
migrate$_1$: A species travels from a region to another region seasonally. migrating$_1$: A species that travels from a region to another region seasonally.			

8.2.1.6 *Translations of collocations*

Equivalence is the topic of Chapter 9, but a specific aspect of equivalence is presented in this chapter since it exploits lexical functions directly.

Lexical functions are language-independent and can be applied to relations and meanings in different languages. Table 8.4 listed some English, French and Spanish collocations encoded with the same LF. More examples are given below: these French and English collocations all contain verbs that express 'creation'.

Caus$_1$Func$_0$(*alias*) = *create, define*	**Caus$_1$Func$_0$**(*alias*) = créer, définir
Caus$_1$Func$_0$(*code*) = *write*	**Caus$_1$Func$_0$**(*code*) = *écrire*
Caus$_1$Func$_0$(*condition*) = *specify*	**Caus$_1$Func$_0$**(*condition*) = *définir, écrire, spécifier*
Caus$_1$Func$_0$(file) = *create*	**Caus$_1$Func$_0$**(*fichier*) = *créer*
Caus$_1$Func$_0$(*folder*) = *create*	**Caus$_1$Func$_0$**(*dossier*) = *créer*
Caus$_1$Func$_0$(*interface*) = *develop*	**Caus$_1$Func$_0$**(*interface*) = développer
Caus$_1$Func$_0$(*page*) = *build*	**Caus$_1$Func$_0$**(*page*) = *construire, créer*
Caus$_1$Func$_0$(*parameter*) = *define*	**Caus$_1$Func$_0$**(*parameter*) = *définir*
Caus$_1$Func$_0$(*password*) = *create, define, set*	**Caus$_1$Func$_0$**(*password*) = *créer, définir*
Caus$_1$Func$_0$(*program*) = *write*	**Caus$_1$Func$_0$**(*program*) = *écrire*
Caus$_1$Func$_0$(*site*) = *construct, create*	**Caus$_1$Func$_0$**(*site*) = construire, créer

If multilingual terminological resources encode relations with ***language-independent*** LFs, translations can thus be accessed easily. More specifically, LFs can be used to locate and retrieve translations of collocations. This section describes a method that exploits this language-independence in the resources DiCoEnviro and DiCoInfo.

The method is designed to allow users to enter an expression in a given language, such as *send a file as an attachment* in English, and retrieve the translations from the other versions of the resources provided that they were previously encoded by terminologists in other languages. In this case, users should be able to retrieve the French equivalent *envoyer un fichier en pièce jointe*. Furthermore, since there can be more than one option in a given language, the method should provide users with all possible translations. For instance, *move a mouse* can be translated into French with *déplacer une souris, manipuler une souris,* and *faire glisser une souris*; similarly *delete a file* can be translated into Spanish as *borrar, eliminar*, and *suprimir un archivo*.

When compilers add entries to the DiCoEnviro or the DiCoInfo, they assign equivalence relationships between headwords (e.g. Fr: *souris* ⇔ En: *mouse* ⇔ Es: *ratón*; En: *Internet* ⇔ Fr: *Internet;* En: *attachment* ⇔ Fr: *pièce jointe*). These equivalents are shown to users in the online interface and hyperlinked so they can access this entry in another language. However, collocations are not translated one by one; they are added separately to each language version.

The general strategy for finding translations for collocations is relatively straightforward. Users enter an expression, such as *move a mouse* in English. In the DiCoInfo, the French *souris* is defined as an equivalent of *mouse. Move* is

listed as a possible collocate in the entry for *mouse* and is encoded with **Real**$_1$ (the verb used to express the typical use of a mouse when carried out by its first argument). We can establish that there is an equivalent collocation in French when a correspondence can be found in the entry for *souris*. In other words, a match is made when collocates encoded with **Real**$_1$ are also found in the *souris* entry. In French, three options are possible as shown below. Interestingly, although the verbs *déplacer*, *manipuler*, and *faire glisser* are encoded with the same lexical function in French with respect to their use within this collocation, they are not true synonyms when considered independently.

	English		French
Entry	*mouse*	⇔	*souris*
Collocate	*move a ~*	**Real**$_1$	*déplacer une ~* *manipuler une ~* *faire glisser une ~*
	click on … with a ~	**Labeal**$_1$	*cliquer sur … avec une ~*

We will examine a slightly more complex example. Here, we wish to retrieve the French equivalent of the English collocation *send an email*. The system resorts to the strategy that was described above with *move a mouse* and looks up the entries to find where this expression was encoded. In this case, the collocation appears in three entries, since *email* is polysemous.

email$_1$, *n.*
Real$_{12}$: *send an ~*

email$_2$, *n.*
Real$_{12}$: *send ~*

email$_3$, *n.*
Labreal$_{12}$: *send … by ~*

Three equivalents for *email* are retrieved from the French version.

email$_1$, *n.* ⇔ *courriel*$_1$, *n. m.* ('an electronic message sent to a recipient')
email$_2$, *n.* ⇔ *courriel*$_2$, *n. m.* ('a series of electronic messages (contained in the inbox, for example)')
email$_3$, *n.* ⇔ *courrier électronique*$_1$, *n. m.* ('application used for sending electronic messages')

The system can then use the lexical functions that appear in the three English entries and try to find a match in the French entries. As was the case with *move a mouse,* there can be more than one expression available in another language for a given English collocation.

	English		French
Entry	*email*$_1$	⇔	*courriel*$_1$
Collocate	*send an ~*	**Real**$_{12}$	*envoyer un ~* *transmettre un ~*
Entry	*email*$_2$	⇔	*courriel*$_2$
Collocate	*send ~*	**Real**$_{12}$	*envoyer du ~* *expédier du ~* *transmettre du ~*
Entry	*email*$_3$	⇔	*courrier électronique*$_1$
Collocate	*send … by ~*	**Labreal**$_{12}$	*envoyer … par ~* *transmettre … par ~*

8.2.2 Semantic frames to discover different kinds of structures

Another method can be used to discover structures in fields of knowledge. It consists of organizing terms in accordance with the cognitive background necessary to understand them. This section will show how the principles of Frame Semantics, FS (Fillmore 1976, 1982; Fillmore and Baker 2010), that was briefly introduced in Section 3.4.2, can be applied to environment terms.

We saw that ***Frame Semantics*** (FS) assumes that the meaning of lexical units is construed on the basis of ***background knowledge***. This background knowledge is captured in ***semantic frames*** which can be viewed as abstract representations of prototypical ***situations***. A situation comprises participants, and other conceptual elements, which constitute its ***frame elements*** (FEs). Finally, in this framework, lexical units are said to 'evoke' semantic frames. LUs in a given frame share the same conceptual components.[87]

Frame Semantics was devised to account for prototypical situations and was mainly applied to "general" language. However, its principles can be extended to specialized situations and to terms that evoke them. Semantic frames should also offer a window on the way situations are conceptualized in special subject fields. However, it must be pointed out that terminologists seldom possess previous 'background knowledge' relative to the domains they are asked to account for, since they are not experts. This knowledge must generally be acquired. The information contained in specialized corpora along with the consultation of experts can assist them in the process.

87. A concrete illustration of a semantic frame designed to capture a 'living somewhere' situation was given in Figure 3.6. Figures 3.7 to 3.10 show how the lexical content of frames is described and how frames can be interlinked.

We will examine how we can apply this framework – and more specifically the modeling defined in FrameNet (Ruppenhofer et al. 2016) – to a small set of terms, more precisely terms linked to *animal* viewed from the perspective of endangered species. The evidence used to support this perspective is supplied by a corpus that contains specialized texts. Consider the examples below:

inhabit, live, occupy, colonize, distribution
Young or small fish are noted to INHABIT *gravel riffles …*
(Mississippi Museum of Natural Science 2014)
The other subspecies of Puma concolor LIVE *in almost any type of habitat … standardize presentation of references (Mississippi Museum of Natural Science 2014)*
The Concho Water Snake OCCUPIES *a restricted geographic range. (Campbell 1995)*
Rare species generally lack an ability to rapidly COLONIZE *areas and they are often poor competitors. (Stohlgren & Sunil 2013)*
The DISTRIBUTION *and abundance of threatened species across forest fragments varied … (Utah State University Extension Service Department of Fisheries and Wildlife 1998)*

live, survive, survival
Whooping Cranes can LIVE *up to 22 to 24 years in the wild. (Campbell 1995)*
fish, salmon, trout and other ocean species that need cold water to SURVIVE *will eventually become extinct. (PANACEA 2015)*
Increase in size of the Sahara may negatively impact SURVIVAL *of palaearctic migratory birds. (Intergovernmental Panel on Climate Change (IPCC) 2006c)*

mate, give birth, produce, breed
Humpback whales inhabiting the waters of the Northern Hemisphere MATE *between October and March. (Benson & Nagel 2004)*
These congregations are large, and the females GIVE BIRTH *to a single young in late May or early June. (Benson & Nagel 2004)*
A female sturgeon may PRODUCE *between 800,000 and 2,500,000 eggs per spawning season. (Mississippi Museum of Natural Science 2014)*
Most of these eagles BREED *in the northern U.S. and Canada. (Utah State University Extension Service Department of Fisheries and Wildlife 1998)*

hatch, grow, mature, fledge
… most eggs HATCH *by the end of May. (Campbell 1995)*
Pallid sturgeon can live up to 60 years, and GROW *to 6 feet (1.8 meters) and 80 pounds (36 kilograms) in size. (U.S. Fish & Wildlife Service 2008)*
Female black bears MATURE *and can have their first litter in three years. (Mississippi Museum of Natural Science 2014)*
Most owlets FLEDGE *(leave the nest) in June, about 35 days after hatching. (Campbell 1995)*

Verbs and nouns that are capitalized in these sentences denote different kinds of activities that affect animals during their lives. In other words, these terms are associated

with different ***situations***: some concern the existence of animals; others denote their being in a given location, other terms are linked to the way they produce offspring; finally, an additional set of terms refer to different stages of life.

Consider the subset of terms relating to animals in given locations. Some terms denote the situation in which animals stay somewhere for a certain time: *inhabit*, *live*, *occupy*. A slightly different situation in which species can be found in a delimited area is evoked by terms such as *distributed*, *distribution*, and *range*. Finally, animals can also settle in a given area in order to feed and reproduce: this situation is expressed by *colonize*, *colonization*, *recolonization*.

All three situations involve two obligatory participants (or core frame elements): the first one corresponds to a living entity (an animal, a species, a bird, etc.); the second one, to a location (a site, a habitat, a range, etc.). Nouns and verbs all convey the basic idea that the living entity finds itself in this location. However, terms provide a slightly different perspective on the general 'being somewhere' situation. Some terms (*inhabit*, *occupy*, and *live*) denote a situation whereby living entities stay in a location where they can carry out various activities. Other terms contain an additional semantic component, that of a delimited area where a given number of animals can be found (*distributed, distribution, range*). Finally, a third set of terms incorporates additional components, i.e. 'settling in an area' and 'being in groups' (*colonization, colonize, recolonization*). We can thus consider that each series of terms evokes a slightly different situation.[88]

We can capture the first situation in a semantic frame similar to the one reproduced in Figure 8.16.[89] A definition of the frame along with examples in English and French are provided in the upper part of the figure. In the definition and

---

88. Of course, there can be some discussion about the way to define situations and delimit frames designed to capture them. Criteria based partly on the linguistic behavior of linguistic units (Ruppenhoffer et al. 2016: 11–17) were devised in order to assist the definition of frames. Among these criteria, we can find:

Same number and types of frame elements (obligatory as well as optional).
Lexical units denote the same part of the scene.
Lexical unit profile the same perspective.

89. This semantic frame is based on a frame called **Residence** in FrameNet (2017). It was adapted to account for the terms in the field of endangered species and to specific methodological choices that can differ from those made in FrameNet. One of them concerns the choice of labels for participants. FrameNet uses very specific labels as it postulates that frame elements are defined within frames; the Framed DiCoEnviro uses a version of semantic roles (see Section 6.5).

examples, the obligatory participants are highlighted and their contribution to the frame is specified. Then, lists of obligatory (Participants 1) and optional participants (participants 2) are listed along with examples for each of them. Finally, English and French terms that evoke this frame are given. It should be mentioned at this point that the contents of frames that are reproduced in this section are by no means complete and reflects preliminary work. Nevertheless, it still gives us an idea of how terminological data can be modeled using Frame Semantics principles.

The two other 'being somewhere' situations could be modeled in new frames. Figure 8.17 shows how the DISTRIBUTION frame is structured. We can readily visualize some similarities but also some differences between this situation and the one described in the RESIDENCE frame. A new obligatory participant is added (**Expanse**) and accounts for the idea that a 'delimited area' is necessary to characterize this situation. Of course, terms that evoke this second frame also differ.

RESIDENCE

Definition: A **Patient** resides in a specific **Location** in order to carry out different activities.

Example(s): En – ***This species*** *does not* INHABIT ***areas with a sandy bottom devoid of vegetation****, nor do they occur where the bottom is muddy, whether or not vegetation is present.*
(Campbell 1995)

Fr – *Les inventaires acoustiques réalisés au cours des dix dernières années au Québec démontrent que* ***la chauve-souris rousse*** OCCUPE ***une grande partie de la province****.*
(Levesque & Tremblay 2008)

Participants 1 (core frame elements)	Participants 2 (non-core frame elements)
Patient: ***Young or small fish*** *are noted to* INHABIT *gravel riffles*	**Condition**: *It can* INHABIT *this area* ***if the temperature is suitable****.*
Location: *The other subspecies* LIVE ***in almost any type of habitat****.*	**Duration:** *these animals* ***still*** OCCUPY *less than 1 percent of their historic habitat*
	Time: *The Black-footed Ferret* ***once*** INHABITED *extensive areas*
	…
English terms:	French terms:
inhabit	*nicher*
live	*nidification*
occupy	*occuper*
	vivre

Figure 8.16 The RESIDENCE frame adapted to environment terms (based on the Framed DicoEnviro 2018)

Distribution

Definition: A **Patient** resides across an **Expanse** or expands from a **Location** to another **Location**.

Example(s): En – *The distribution and abundance* ***of threatened species across forest fragments*** *varied markedly (Figure 3).*
(Utah State University Extension Service Department of Fisheries and Wildlife 1998)
Fr – *La connaissance actuelle de la répartition et de l'abondance* ***des grands singes dans l'Est de la RDC*** *se limite à une superficie relativement restreinte où les scientifiques ont pu travailler durant quasiment 20 ans de troubles politiques.* *(UICN 2012)*

Participants 1 (core frame elements)	Participants 2 (non-core frame elements)
Patient	**Degree**
Expanse	**Frequency**
Location	**Manner**
Location1	**Purpose**
Location2	**Time**
English terms:	French terms:
distributed	*distribution*
distribution	*réparti*
range	*répartition*

Figure 8.17 The Distribution frame adapted to environment terms (based on the Framed DiCoEnviro 2018)

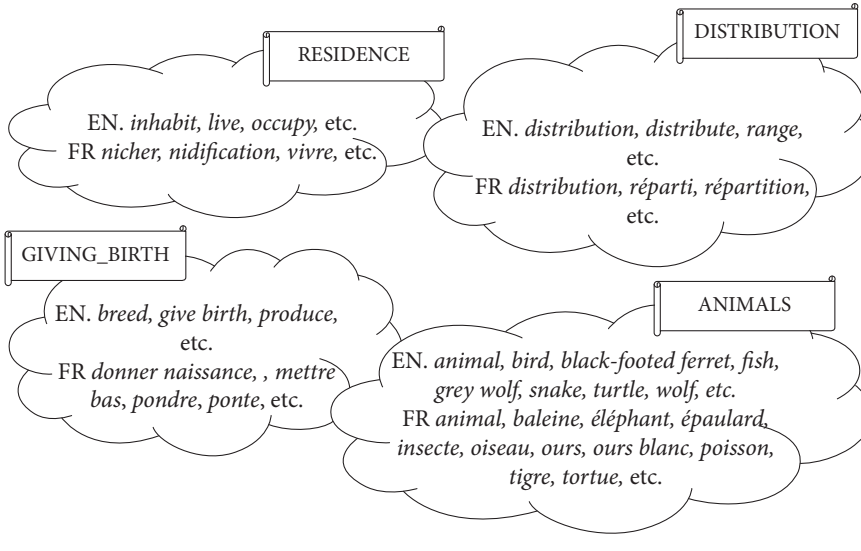


Figure 8.18 Frames for terms related to endangered species

We can go on like this and define new frames to capture other situations that concern animals (those that evoke reproduction; those related to going through development

stages, etc.) as shown in Figure 8.18. Frames can also be defined for terms that denote entities and not only those that denote activities. This way we obtain a first picture of how terms connect to given situations in a field of knowledge. However, this view remains partial. The next section explains how frame modeling can further help us obtain a much broader picture of important situations in a given domain.

8.2.2.1 *Obtaining a better view of related situations with frames*

In the previous section, we mentioned that some situations – even if they are modeled in separate frames – are closely related. The situations captured in the RESIDENCE and DISTRIBUTION frames are more closely connected than other situations in which species are involved such as reproducing, surviving and eating.

In order to account for these relationships, semantic frames can be connected via different sets of links. Figure 8.19 gives a broader perspective on 'being somewhere' situations applied to species. Our RESIDENCE frame is connected to a frame called DISTRIBUTION evoked by the terms *distribution*, *range* and *distributed*. RESIDENCE is also linked to the frame SPECIES_COLONIZATION that contains the terms *colonize, colonization* and *recolonization*. We can then connect SPECIES_COLONIZATION to another frame describing a situation in which an **Agent** places animals in a given location. This new frame is evoked in English by terms such as *introduce*, *introduction*, *reintroduce* and *reintroduction*.

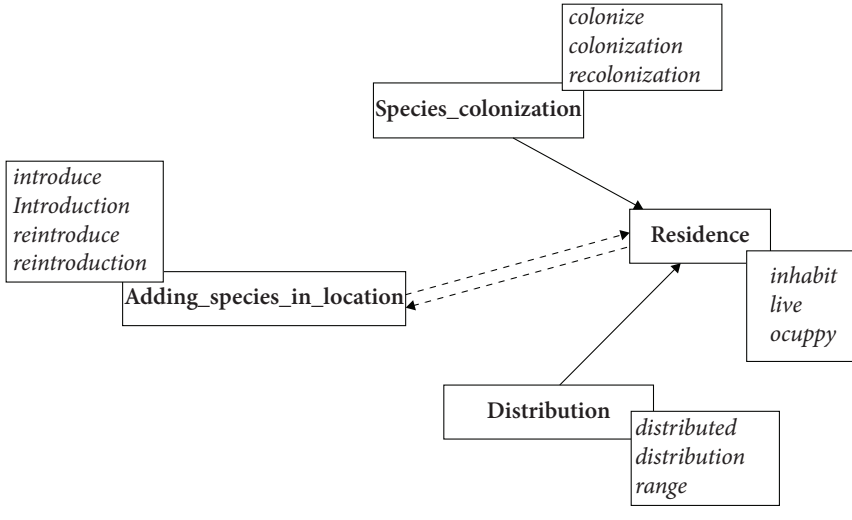


Figure 8.19 "Being somewhere" situations for species (based on the Framed DicoEnviro 2018)

Of course, species carry out many other activities. They are born, then they grow, they can reproduce and they die. They also migrate and feed on different kinds of food.

Finally, they can be endangered by different threats, survive these threats or become extinct. We can model all these situations and connect them via meaningful links.

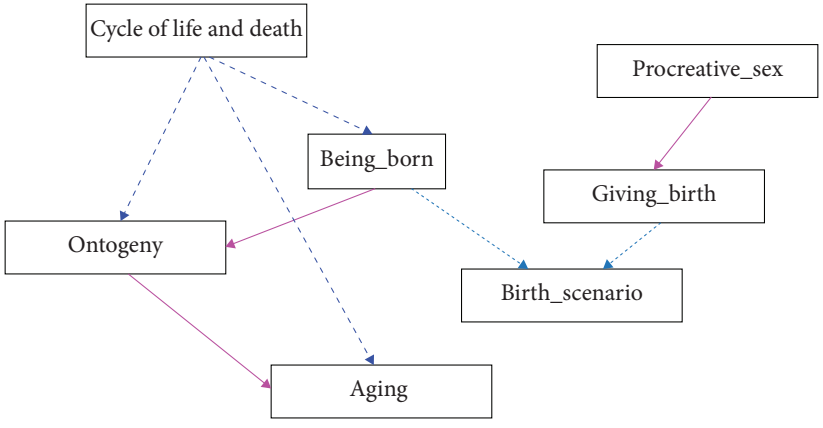


Figure 8.20 Situations related to the cycle of life and death of endangered species (based on the Frame DiCoEnviro 2018)

Firstly, life cycle (birth, growth, reproduction and death) occurs in a chronological sequence for which frames can account. This is shown in Figure 8.20 in which the different life stages of species are defined as subframes in a general CYCLE_OF_LIFE_AND_DEATH scenario. Subframes relations are represented with dotted blue arrows.

Secondly, some situations occur before other ones. Before they give birth to offspring, species first mate; before they mature, they grow. This is represented in Figure 8.20 in which a precedence relation was established between the frames PROCREATIVE_SEX and GIVING_BIRTH and between the ONTOGENY and AGING frames (pink arrows).

Thirdly, general situations can be conceptualized from two different perspectives. For instance, birth can be considered from the point of view of the parent giving birth. This situation is modeled in the GIVING_BIRTH frame which foregrounds the parent and places the offspring in the background (this situation is evoked by terms such as *give birth* and *produce* in English and *donner naissance*, *mettre bas* and *mise bas* in French). The situation can also be considered from the perspective of the offspring that comes to life. The BEING_BORN frame accounts for this perspective by foregrounding the offspring and placing the parent in the background (this frame is evoked by terms such as *born* and *hatch* in English and *éclore*, *naissance* and *naître* in French). Figure 8.20 shows how the relation between the two frames is captured through their connection to a more general BIRTH_SCENARIO frame for which they provide a different perspective (light blue arrows).

8.2.2.2 *Highlighting differences between specialized and general knowledge*

Fillmore (1982) and other authors have pointed out that some situations are conceptualized differently in everyday and specialized contexts (see for instance the example with 'innocent' and 'guilty' given by Fillmore, Section 5.2.5). Indeed, the background knowledge that applies to specialized situations is likely to differ from everyday knowledge since it must be acquired through formal training. Frame Semantics should allow us to understand this difference and perhaps better characterize it.

We will use an example that was mentioned in previous chapters, the verb *warm,* and show how we can spot differences. In Section 5.2.5, we explained that the verb undergoes ***meaning modulations*** when considered from the point of view of climate change. A general description such as the one provided in FrameNet covers a very broad spectrum, much of which does not apply to the uses of the verb in this specialized domain. We explore the matter further in this section and compare the modeling proposed in FrameNet and the one given in the environmental resource Framed DiCoEnviro. In this example, it is assumed that FrameNet models everyday situations. The comparison is summarized in Table 8.6.[90]

In FrameNet (2018), *warm* appears in the frame CAUSE_TEMPERATURE_CHANGE that is defined as a situation in which 'An **Agent** changes the temperature of an **Item**. A **Temperature_goal** can specify the desired temperature. A **Temperature_change** can also be indicated. The **Temperature_start** indicates the initial temperature.' (FrameNet 2018) This situation includes four core frame elements: **Agent**, **Cause, Hot_cold_source**, and **Item**. It can also involve non-core frame elements, such as **Circumstances**, **Manner** and **Purpose**. In the Framed DiCoEnviro, the situation described as CAUSE_TEMPERATURE_CHANGE, and conceptualized from the perspective of climate change, includes two obligatory participants, i.e. **Cause** and **Patient**,[91] and optional participants such as **Degree**, **Duration** and **Result**.

90. It should be kept in mind that the data in the Framed DiCoEnviro column is incomplete, since the resource in under construction. Nevertheless, some important differences between the frame described in FrameNet and the one given in the Framed DiCoEnviro can still be observed.

91. Although labels in FrameNet and the Framed DiCoEnviro differ (see Chapter 6, in which labeling systems are explained), there is no fundamental difference between **Item** in FrameNet and **Patient** in the Framed DiCoEnviro.

Table 8.6 Characterization of the frame CAUSE_TEMPERATURE_CHANGE in FrameNet and the Framed DiCoENviro

CAUSE_TEMPERATURE_CHANGE (FrameNet 2018)	CAUSE_TEMPERATURE_CHANGE (Framed DiCoEnviro 2018)
Definition	
An **AGENT** changes the temperature of an **ITEM**. A **TEMPERATURE_GOAL** can specify the desired temperature. A **TEMPERATURE_CHANGE** can also be indicated. The **TEMPERATURE_START** indicates the initial temperature.	A **CAUSE** changes the temperature of a **PATIENT**.
Core frame elements / Obligatory participants	
AGENT (***Ryan*** *reheated the pasta in the microwave*)	
CAUSE (***The Sun*** *warmed the house*)	**CAUSE** (***An increase in greenhouse gases*** *warms the atmosphere*)
HOT_COLD_SOURCE (*We chilled the drink* ***on ice***)	
ITEM (*Ryan reheated* ***the pasta*** *in the microwave*)	**PATIENT** (*Positive radiative forcings warm* ***the Earth 's surface***)
Non-core frame elements / Optional participants	
CIRCUMSTANCES (*Bill cooled the Jello* ***at room temperature***)	
CONTAINER (*Heat the potatoes* ***in a medium-sized pan***)	
DEGREE (*You will probably have to wait a while before you warm your legs up* ***completely***)	**DEGREE** (*Water vapour warms the atmosphere* ***significantly***).
DURATION (*Refrigerate the cookie dough* ***for an hour*** *before cooking*)	**DURATION** (*cooling the Earth by about 0.5 degrees C* ***for up to a year***)
	EXPANSE (*we are currently concerned about* ***global*** *warming*)
INSTRUMENT (*Ryan reheated the pasta* ***in the microwave***)	
MANNER (*Frozen solid in winter, they WARM* ***quickly*** *in spring sunshine*)	
MEANS (*Heat the beeswax and almond oil* ***in a double enamel boiler***)	
PLACE (**In the laboratory** their eggs can be heated to 98°C)	**LOCATION** (*… warming temperatures* ***at the poles***)
PURPOSE (*The saline is eventually filtered and then HEATED in shallow pans* ***to complete the evaporation process before the salt is left to crystallise***)	
RESULT (*The Sun itself is destructive*, heating *the rocks by day* ***so they expand***)	**RESULT** (*The aerosol direct forcing cools the surface,* ***altering atmospheric stability***)

Table 8.6 (*Continued*)

Cause_Temperature_Change (FrameNet 2018)	Cause_Temperature_Change (Framed DiCoEnviro 2018)
SUBREGION (*The mousse was chilled **around the edges***)	
TEMPERATURE_CHANGE	
TEMPERATURE_GOAL (*chilling the cotyledonary petiole **to 0.5 °C** extends the time for spontaneous recovery to about 25 minutes*)	**VALUE** (*cooling the Earth **by about 0.5 degrees C** for up to a year*)
TEMPERATURE_START	
TIME (*It is best to chill the beer slightly **before drinking***)	**TIME** (*the estimated net effect of these perturbations is to have warmed the global climate **since 1750***)
Lexical units / Terms	
chill.v, cool down.v, cool.v, heat up.v, heat.v, overheat.v, refrigerate.v, reheat.v, warm up.v, warm.v	$cool_{1b}$, $cooling_{1b}$ $warm_{1b}$, $warming_{1b}$

Both resources describe a situation in which something causes something else to change temperature. However, the content of the Framed DiCoEnviro differs significantly from that of FrameNet. First, the environment frame does not include an **Agent**: only **Causes** can make temperature change (e.g. *the greenhouse effect, perturbations, carbon dioxide*). It does not include a **Hot_cold_source** either, which would be different from the actual **Cause**. Other differences can be seen in the types of optional participants that can characterize the "general" situation as opposed to the environmental one. Finally, the linguistic expressions that are used to realize participants linguistically as well as the lexical units that can evoke each frame differ quite drastically.

In addition to differences between the description of frames and their lexical content, relations established between frames differ depending on the resource. There are various ways in which changes can occur in the environment and they can affect different kinds of natural entities. There are also multiple terms that are used to evoke these situations (*change*, *warm*, *melt*, *retreat*, etc.).

8.2.2.3 *Capturing meaning modulations and different conceptualizations within the same domain*

In addition to highlighting differences between general language and specialized domains, meaning modulations can have an impact on meaning distinctions within the same domain. Frame Semantics explains this phenomenon by the fact that situations are conceptualized differently and attempts to account for these differences in specific frames.

Consider an example that was given in Chapter 5 (Section 5.2.5): the verb *hunt* and distinctions that appear to be valid only if the verb is considered from the point of view of endangered species. A hunting situation is considered differently when it is carried out by humans and when it is carried out by other meat eaters. Linguistic as well as conceptual evidence supports this distinction. We could then hypothesize that hunting is conceptualized differently when it is carried out for survival and when it is carried out by human beings for other reasons.

Frame Semantics offers an interesting reading of this phenomenon. The situations would be described in two different frames: we will call them HUNTING and HUMAN_HUNTING respectively. The first one (HUNTING) describes a situation in which a carnivore (a predator) accomplishes a series of activities from chasing to killing to capture another species (a prey) in order to feed itself or its community. This is a necessary activity for survival. This frame can be evoked by the verbs *prey* and *hunt* and by the nouns *hunting* and *predation*. The second frame (HUMAN_HUNTING) models a situation where a human being (a hunter, a poacher) accomplishes a series of tasks from chasing to killing in order to capture an animal (a prey). Usually, these activities are carried out with the help of instruments (a gun, a knife, a crossbow). Moreover, this kind of hunting can be done for different reasons: food, clothing, or even leisure, and can be detrimental to animals if it is not controlled. This frame is evoked by verbs such as *hunt* and *poach* and by nouns such as *hunting* and *poaching*. Figure 8.21 illustrates the differences between the two situations.

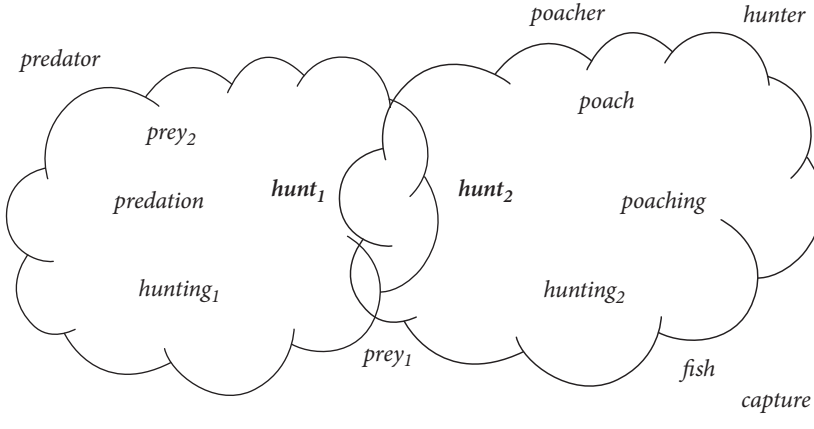


Figure 8.21 *Hunt*$_1$, *hunt*$_2$ and related terms

Considering a broader spectrum of situations in a domain such as the environment helps identify many other differences in conceptualization. For instance,

since the HUNTING and HUMAN_HUNTING were defined as two different situations, they appear in a different network of related situations. The HUNTING situation, as it concerns the survival of species, is linked to other activities carried by these species. Similarly, the HUMAN_HUNTING situation is linked to activities carried out by men, most of which pose a threat to species and the environment in general.

Figure 8.22 shows how part of the activities carried out by species are captured in frames. The HUNTING frame appears alongside frames that account for other activities in which species are involved: feeding (INGESTION), being somewhere (RESIDENCE), moving from one place to another (SELF_MOTION), reproduction (PROCREATION), etc. Figure 8.23 accounts for activities carried out by human beings also modeled in semantic frames. The HUMAN_HUNTING frame is connected to the frame USING_RESOURCE with other frames (AGRICULTURE, MINING and FISHING). The differences between the two sets of situations are striking. Based on this modeling, we can say that species are conceptualized in the field of the environment as living creatures struggling to survive. In contrast, human beings are conceptualized as users (and often over-users) of natural resources.

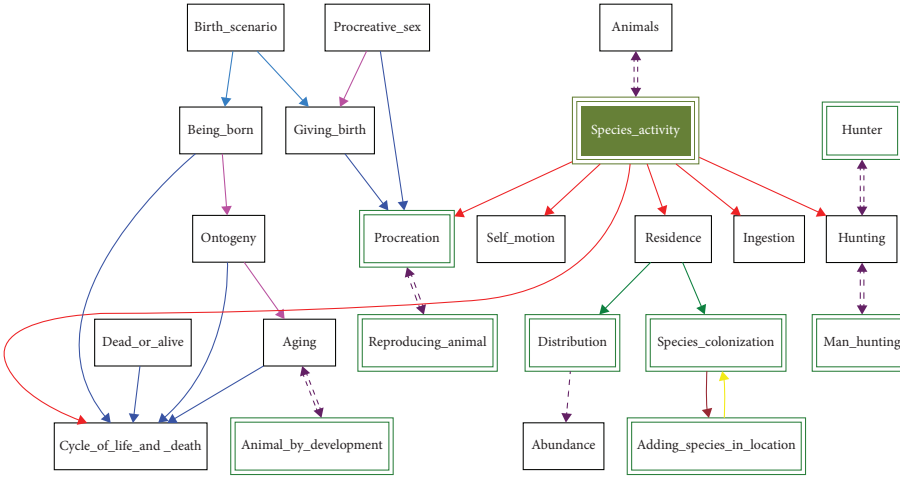


Figure 8.22 A sample of activities carried out by species (based on the Framed DicoEnviro 2018)

Many more examples could be given to show how the principles of Frame Semantics and the methodology devised in FrameNet can be used to get a better understanding of specialized situations and how we conceptualize them. The work on the application of Frame Semantics to terminology is relatively recent, but it is very promising.

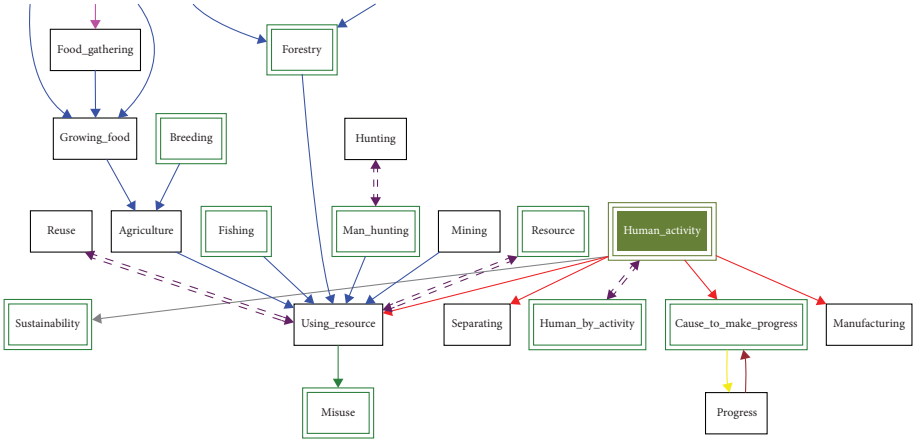


Figure 8.23 A sample of activities carried out by human beings (based on the Framed DicoEnviro 2018)

Summary

This chapter introduced different ways of constructing and exploiting conceptual and terminological structures.

Conceptual structures account for the organization of knowledge. Since they are based on concepts, a formal distinction is made between concepts and designations: designations, when added, appear in a separate module that can be appended to the core conceptual system. Conceptual structures can be based on a single set of relations (such as taxonomic or partitive ones) or combine multiple relations. Tree representations or graphs can be used to represent structures graphically. Different kinds of operations can be carried out on large sets of concepts once they are formally defined in a conceptual system: inheritance (single or multiple), definition of disjoint categories, addition of new relations, etc.

Terminological structures usually account for various kinds of relations between terms. Paradigmatic and syntagmatic relations can be captured with the system of lexical functions. LFs are used to encode relations in different languages to reveal the terminological structure of a domain in different ways. Another way of understanding the terminological structure of a domain consists in representing the situations denoted by terms in semantic frames and linking these frames.

With the exception of a short subsection, Chapters 7 and 8 considered relations from the point of view of a single language. The next chapter examines relations established across languages and some of the challenges that they raise even in conceptual approaches where concepts are implicitly assumed to be language-independent.

Further reading

A nice way to understand formal conceptual structures is to learn how to design ontologies. An excellent starting point is the tutorial by Foy et al. (2014). We already mentioned EcoLexicon (2018) at the end of the previous chapter. It is also suggested here to explore the complex modeling of the conceptual structure of the environment. EcoLexicon also implements nice features to generate different kinds of views on the relations.

Models for definitions based on knowledge-based approaches are dealt with in León Araúz et al. (2012) and Seppälä et al. (2016).

WordNet is extremely useful to understand how the structure of the lexicon can be represented in a lexical resource with a large coverage. Fellbaum (1998) contains contributions that describe the general principles on which WordNet is based. Another lexical resource that shows multiple relations between lexical units in many different languages, terms and encyclopedic data is BabelNet (2017; Navigli and Ponzetto 2012).

A collective book edited by Wanner (1996) contains interesting contributions on lexical functions. Readers who can read French can also consult Meľčuk et al. (1995). The Réseau lexical du français (Polguère 2014) is a sophisticated resource that encodes paradigmatic and syntagmatic relations between lexical units with lexical functions.

Frawley (1988) and L'Homme (2002) argue for the use of lexical functions to represent terminological relations. Two terminological resources encode relations between terms with lexical functions: the DiCoEnviro (2018) and the DiCoInfo (2018). A graph modeling of relations based on lexical functions is proposed in the DiCoInfo visuel (Robichaud 2012) and the NeoVisual (L'Homme et al. 2018).

In L'Homme et al. (2015), the different methods for presenting collocations in the DiCoInfo and for retrieving translations are described. San Martín and L'Homme (2014) explain the rules applied in the DiCoEnviro and the DiCoInfo for writing definitions.

A specialized resource using Frame Semantics describes situations with respect to soccer, i.e. the Kicktionary (Schmidt 2009). L'Homme (2018) argues for the use of Frame Semantics to structure knowledge in the field of the environment.

CHAPTER 9

Equivalence in terminology

This chapter addresses the issue of ***relations*** across languages, more specifically relations of ***equivalence***, and some of the challenges that they raise. Up to this point we have been concerned with terms from the point of view of a single language. Cross-linguistic matters were only mentioned occasionally. In practice, however, terminology work is often carried out in a bilingual or a multilingual setting and most terminological resources document terms and related data in more than one language.

This chapter is much shorter than the previous ones on conceptual and terminological relations and structures. This should not be interpreted as the small importance given to equivalence. Rather, it seems that equivalence has not been debated in terminology as much as in other fields. Sager pointed out the following a few decades ago:

> In terminology we find little, if any reference to the nature of translation equivalents, because the theory of monosemic reference of term to concepts does not readily admit to problems of equivalence. (Sager 1994: 55)

It seems fairly obvious that dealing with terms that designate specialized realities is not as problematic as trying to establish equivalence in other situations, for instance, when important cultural specificities are involved. Furthermore, the many efforts made by scientific communities towards finding ways to improve communication and the international nature of research and development probably contribute to reduce the complexity of transferring knowledge from one language to another. That being said, even in terminology, establishing equivalence between two terms or concepts is not always a straightforward task as it may seem at first. Section 9.3 examines different problems posed by equivalence relations.

In the next two sections, the distinction made in previous chapters between knowledge-driven and lexicon-driven perspectives will be maintained. First, equivalence is introduced from the point of view of knowledge-based approaches and will be called ***conceptual equivalence***. Then, it is presented according to a lexicon-based perspective and will be labeled ***terminological equivalence***. Again, this division may appear artificial at times since terminologists probably mix both

approaches when establishing equivalence relations. However, it helps better stress differences as it has in previous chapters.

9.1 Conceptual equivalence

Knowledge-driven approaches to terminology consider that terms are equivalent if they belong to different languages and denote the same concept within the same domain. According to these requirements, equivalents should be listed on the same term record in a term bank or in the same entry in another kind of terminological resource. Examples of this situation were given in Chapter 2. Figure 2.4 shows that English and French terms denoting 'global warming' appear on the same term record in TERMIUM Plus®: En. *global warming, global-warming*; Fr. *réchauffement climatique, réchauffement planétaire, réchauffement de la planète, réchauffement général de la planète, réchauffement du globe, réchauffement du Globe, réchauffement global*. Figure 2.5 shows that terms that denote the concept 'habitat' in several languages are all listed in a single entry: Ar. موطن; Eu. habitat; Cs. stanoviště; Zh. 栖息, and so on.

Theoretically, knowledge-driven approaches consider that concepts are language-independent entities, so the issue of establishing an equivalence relation should boil down to finding the right designations for a clearly delineated concept and to do this in different languages. In other words, the focus is on ***conceptual equivalence***. Knowledge-driven approaches are concerned first and foremost with ***exact equivalence***, a relation in which terms that belong to different languages denote a single concept. They should label the same node in a conceptual structure. We can illustrate this by reproducing a conceptual structure that was already given for exact synonymy (Figure 7.4). In this version (Figure 9.1), new labels in French were added and distinctions were made between English (En) and French (Fr) designations. All the labels that appear in Figure 9.1 designate the chosen concept. In this particular case, there are multiple designations in English and in French. Designations in the same language are called *exact synonyms*; designations in different languages, ***exact equivalents***. Theoretically, we could even add equivalents in other languages without altering the structure itself.

When establishing equivalence, terms are considered from the point of view of their function that consists in labeling a given concept. Looking back on the *global warming* example, we saw that TERMIUM Plus® suggests a long list of French equivalents: *réchauffement climatique, réchauffement planétaire, réchauffement de la planète, réchauffement général de la planète, réchauffement du globe, réchauffement du Globe, réchauffement global*. These might all be valid equivalents for the concept described, at least they were deemed as such by the compiler of the record.

However, the internal structures of French terms emphasize a particular aspect of the concept. *Réchauffement climatique* stresses the fact that the warming affects the climate; *réchauffement planétaire, réchauffement de la planète, réchauffement du globe,* and *réchauffement du Globe* focus on the fact that the warming affects the planet; *réchauffement global* and *réchauffement général de la planète* convey the idea of a generalization of the warming.

Knowledge-driven perspectives have given little importance to the structure of terms when establishing equivalence. This is why equivalents can correspond to lexical units in the sense defined in Section 4.3, but also to collocations or compositional multi-word terms. Different types of linguistic units can be defined as equivalents and presented as such on term records.

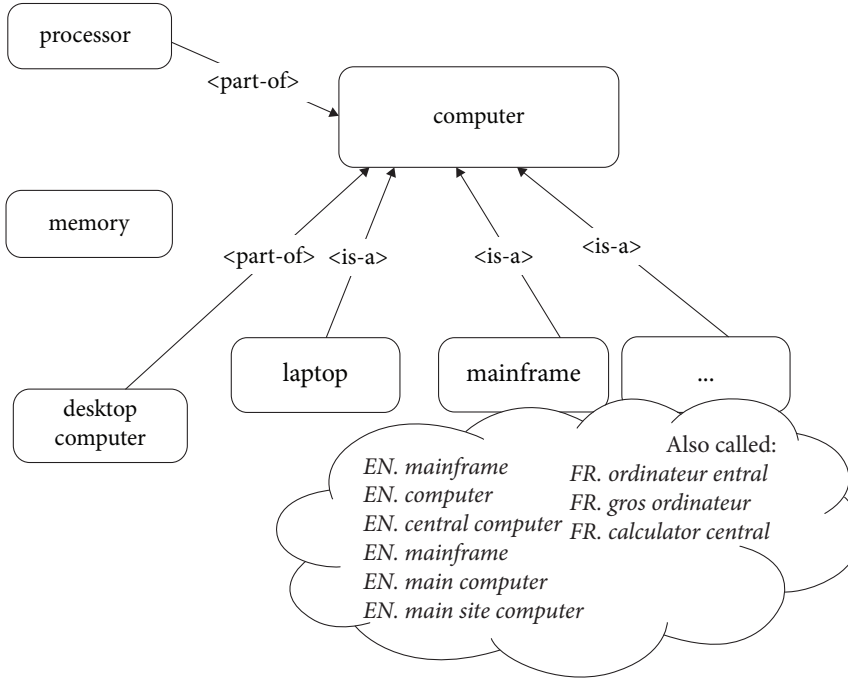


Figure 9.1 Labels in English and French in a conceptual structure

9.2 Terminological equivalence

Equivalence can also be established between terms on the basis of the meaning they carry rather than according to their potential to label a predefined concept. This is the perspective taken by lexicon-driven approaches. ***Terminological***

equivalence is defined as the relation between terms that belong to different languages and that convey the same meaning in the same domain. For instance, En. *ecosystem*; Fr. *écosystème*; Es. *ecosistema* are equivalents in the environment, since their meaning can roughly be paraphrased as follows: 'an ecological unit composed of living things and the conditions in which they live that functions as a whole'.

If terms in different languages convey exactly the same meaning, as *ecosystem*, *écosystème*, *ecosistema*, then this ***equivalence*** is defined as ***exact*** as in knowledge-based approaches. We said on multiple occasions that lexical items can be polysemous. This inevitably affects equivalence. Hence, in lexicon-based approaches, exact equivalence is defined between two lexical units and not between two lexical items.

Based on what was just said, three different situations can occur when considering lexical items in two different languages:

- A lexical item carries (at least) two different meanings: the first one is associated with general language; the second one, with a specialized domain. For example, the noun *key* is used in general language where it designates 'a small object used to open doors or boxes. In computing, it designates 'a part of a keyboard pressed by a user to insert a character or send a command' (among others). When it applies to a small object, it translates into French as *clé*. However, when it designates the part of a keyboard, it translates as *touche*.
- A lexical item conveys (at least) two different meanings: these meanings are connected to different fields of knowledge. The noun *dump* is used in waste management and in computing. In waste management, it is defined as 'a specific place where waste is placed' and translates into *dépotoir* or into *décharge* in French;[92] in computing, it designates 'an operation that consists in empying the memory' and its French equivalent would be *vidage*.
- A lexical item carries (at least) two different meanings: these meanings coexist in the same domain. The example of the French *terre* used in the domain of the environment (and mentioned in Chapter 3) will serve to illustrate this situation. We already established that *terre* has four different meanings:[93]

92. *Dépotoir* complies with usage in Québec; *décharge* is more common in France.

93. One might argue that the first *Terre* does not belong to the same lexical item than the other three. This might be the case, but the other three meanings are sufficient evidence to support the point I wish to make here.

Terre$_1$, n. f. 'a planet of the solar system inhabited by living organisms.' (*la* TERRE *est le cadre de variations climatiques*)
terre$_2$, n. f. 'a surface of the Earth occupied by continents or islands and not covered with water'. (*Expression des forces naturelles qui agitent les éléments – la* TERRE*, la mer, l'atmosphère*)
terre$_3$, n. f. 'an area of ground used for specific purposes', e.g. farming. (*conversion de* TERRES *forestières en terres agricoles*)
terre$_4$, n. f. 'the substance in which plants grow'. (*couche de* TERRE *végétale*)
The first *Terre* translates into English as *Earth*; the second as *land*, the third also as *land*, and the fourth as *earth* or *soil*.

A second example is borrowed from Van Campenhoudt (1996:284). The term *watch* in the maritime domain translates into French as *veille* when it denotes the 'action of watching'; into *quart* when it denotes 'a portion of time when a ship's crew is on duty'; and into *bordée* when it designates 'a crew on duty'.

Separate meanings attached to the same lexical item in one language are not always expressed with different lexical items in another, as lexical items can be polysemous in the second language as well. For instance, *mouse* 'a small rodent' and *mouse* 'an input device' are both translated into French by *souris*. But technically, we should consider that we are dealing with two lexical units in English – *mouse*$_1$ and *mouse*$_2$ – and two separate equivalents in French – *souris*$_1$ and *souris*$_2$. In other languages, these meanings could correspond to different lexical items. *Mouse*$_1$ in Spanish translates into *ratón*, but some Spanish-speaking countries use *mouse* to render *mouse*$_2$. Another example is the item *cloud* that refers to 'the visible mass of vapor that floats in the sky', but also to 'a series of online resources and services'. It translates in both cases into French as *nuage*. A closer look at the previous example with *terre* reveals that two of its meanings can be translated into English with *land*. So *land* as well is polysemous, though its different meanings do not parallel those of *terre*.

Section 9.1 briefly explained how knowledge-driven perspectives represent equivalents in a conceptual structure. Figure 9.1 shows how the English and French designations are all attached to the concept 'mainframe'. It can be surmised from this figure that the other concepts in the structure could be associated with one or several designations in each language. We could also add designations in other languages (Spanish, Chinese, Italian, and so on). If no designation is available in a given language, various strategies can be used to fill the gap. This will be explained in Section 9.3.

In lexicon-driven approaches, cross-linguistic analyses are likely to lead to situations where equivalence does not present such an orderly picture. Figure 9.2 shows what can be obtained when comparing some of the meanings of a set of

polysemous lexical items starting with the four meanings of *terre* mentioned earlier. Some of these meanings lead to polysemous English equivalents that can be further associated with French polysemous equivalents and so on. In Figure 9.2, lexical items are indicated in red, and lexical units are presented with a sense number. In addition, a short description of each meaning is provided between quotation marks (e.g. 'A planet of …'). Equivalents are found when two lexical units in English and in French are connected to the same meaning. The different connections produce a very complex network of equivalence relationships where meaning distinctions hardly overlap in English and French. It should be pointed out that only part of the meanings attached to the lexical items are considered here. So the real situation is even more complex.

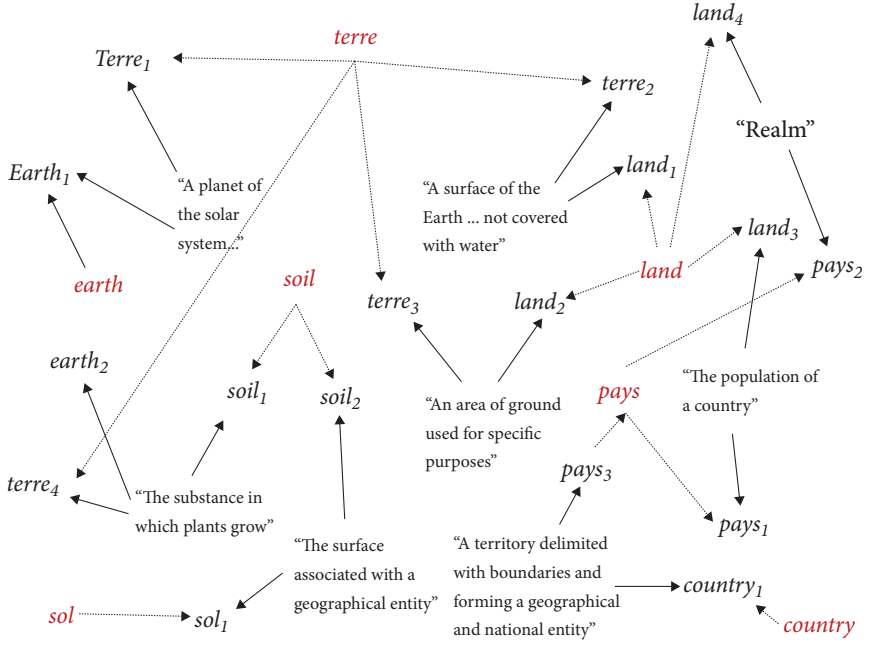


Figure 9.2 Polysemous items and cross-linguistic relationships

9.3 Problems when establishing equivalence

In principle, all languages should be equipped with means to express all sorts of meanings. However, in practice, different situations can prevent the establishment of exact equivalence between two languages. The following subsections describe some of these situations.

9.3.1 Non-equivalence

There are cases where language B lacks an adequate equivalent to express the meaning conveyed by a term in language A resulting in ***non-equivalence***.

This situation first occurs when the terminological apparatus accompanying a new technology or a discovery was developed in the linguistic community in which these novelties appeared. Other cultures might need time to adjust, adapt foreign terms or create equivalents in their own languages. In the meantime, different strategies can be used to fill the gap: direct borrowing with or without an explanation in the native language, literal translation, paraphrase, adaptation or creation of a new designation. Translators are quite familiar with these strategies. A classic example of this situation is information technology terminology where designations first appear in English.

En. *WYSIWYG*
    Fr. *WYSIWYG (ce qu'on voit à l'écran est ce que l'on obtient à l'impression*) (explanatory paraphrase)
En *hashtag*;
    Fr *hashtag* (direct borrowing)
    mot-clic (creation of a new designation)
En *deep learning*
    Fr *apprentissage profond* (literal translation)

In some circumstances, the proposal of equivalents can be carefully engineered. There are scientific communities that established rules to name concepts in different languages with which experts must comply. The adaptation of terms from other languages can also be a part of the language planning activities of states. Official organizations are responsible for reacting to borrowed terms and defining strategies to manage their use amd/or replacement. This is the case in the province of Québec where the *Office Québécois de la langue française* plays that role at an official level. The organization defined a set of criteria and rules for the use and possible adaptation of borrowed terms into Quebec French (Office Québécois de la langue française 2017). When an equivalent becomes available in the native language as a result of language planning activities, we thus obtain an exact equivalent by default. This equivalent is created specifically to designate a predefined concept.

A different kind of situation can lead to non-equivalence. In some cases, concepts themselves are linked to a conceptual system defined within a specific community for which there is no real correspondence in another community. This situation differs from the previous one in the sense that concepts themselves exist in community A, but not in community B. The field of law is often used to illustrate this situation. The entire set of concepts related to Common Law were created in an English-speaking community and therefore labeled in English.

Other linguistic communities use different legal systems, for instance Civil law. However, there is no straightforward correspondence between concepts of Civil Law and Common Law: Designations in one system cannot be used as equivalents for another system.

This situation is a real challenge in Canada where the legal system of most provinces (except Quebec) is based on Common Law. When the Province of New Brunswick became officially bilingual, there was an urgent need for defining French equivalents for everything regarding the legal infrastructure of the province. A complete French terminology was created to translate English terms (Snow 1989, 2010) and made available to users of this terminology.

9.3.2 Partial equivalence

In other situations, the meanings of terms in languages A and B do not overlap perfectly giving way to what has been called ***partial equivalence*** (or ***animorphism***).

A common case of partial equivalence occurs when a language makes a distinction that is not made in the other. A classic example is that of types of wood that Spanish distinguishes as *leña* 'wood used for heating' and *madera* 'wood used for construction'. No lexical equivalent in French or English allows us to maintain this distinction and making it explicit in one of these two languages would require additional explanations. Another example is that of the verb *to farm* that denotes an activity that consists in raising animals or cultivating land for food. In French, no verb covers the same broad meaning. The verb *élever* denotes an activity that consists in raising animals and *cultiver*, an activity that consists in growing and harvesting vegetables or plants. Hence, *to farm* covers both *élever* and *cultiver*.

Of course, partial equivalence always depends on the pair of languages taken into consideration. Looking back on the example of *leña* and *madera*, there is exact equivalence between Portuguese and Spanish since both languages make the same distinction. However, there is partial equivalence between English and Spanish and between French and Spanish.

Interestingly, some strategies for filling gaps generated by non-equivalence have also been used for partial equivalence. For instance, the activity that consists in transferring a file from a distant computer to a local one is designated in French by the verb *télécharger*. In English, a distinction is made depending on the direction of the transfer: *download* is used if the file is placed on the user's computer; in contrast, if the file is sent to a remote computer, *upload* applies. Even if the verb *téléverser* was created in French to render *upload* and parallel the distinction made in English between *download* (télécharger) and *upload* (téléverser), speakers of French often use *télécharger* where *téléverser* would apply.

The examples given above illustrate a particular case of partial equivalence where the meaning of a term in language A is much more general than that of the terms in language B. In fact, the meaning of the term in language A includes the meanings of both terms in language B. Another form of partial equivalence occurs when the meanings of terms in different languages only partly overlap.

9.3.3 Structural divergences

Examples given so far show that equivalents (conceptual and terminological alike) are terms that belong to the same part of speech. This additional condition can raise some difficulties when establishing equivalence relations even between closely related languages.

In some cases, the meaning expressed by a term that belongs to a given part of speech has no direct equivalent in another language. For instance, *bookmark* can be used as a verb or a noun in English (*When you "BOOKMARK" a site it is added to a list your browser saves for you; add a BOOKMARK to your current search*). In French, although a noun translates the English noun *bookmark*, i.e. *signet* (*Dans Explorer, ces SIGNETS se nomment Favoris*); the English verb must be translated by a collocation (*METTRE un site EN SIGNET*). Another example can be given with the pair *s'éteindre: extinction* in French when it applies to endangered species. In English, the noun translates into *extinction* (*Earth's species are already at risk of EXTINCTION*); the verb however, is rendered with the collocation *become extinct* (*Many species of animals have BECOME EXTINCT at least locally*).

There might even be gaps in one of the languages under consideration. For instance, *sustainable* and *unsustainable* translate into French as *durable* and *non durable*. However, no term corresponds directly to *sustainability* and *unsustainability*. A paraphrase such as "le caractère durable de …" will need to be used.[94]

A common equivalence problem arises when terms that fill a certain syntactic function have equivalents that belong to different parts of speech. For instance, *computer* normally translates into French as *ordinateur*. However, when it is used as a modifier of a noun – as in *COMPUTER file* – it translates into *informatique* (*fichier INFORMATIQUE* and not ?*fichier d'ordinateur*). This is a common situation between English and Romance languages. The latter tends to use relational adjectives (see Section 7.2.2.4), whereas English can easily modify nouns with other nouns.

Other forms of ***structural differences*** specifically affect predicative terms.

---

94. In fact, *durabilité* can be found in French texts on environment, but it is much less common than *sustainability* in English. In French, the noun might have been created to translate the English more directly.

- Structural differences between languages occur when a term in language A has a different syntactic behavior than a term in language B. For instance, the verb *click* in English can be used transitively or intransitively (*CLICK an object; CLICK on an object*); in French, the equivalent *cliquer* can only be used intransitively (*CLIQUER sur l'icône*, but not **CLIQUER l'icône*).
- Some predicative terms that are considered to be equivalents do not have the same number of arguments in languages A and B. For example, *Le grand dictionnaire terminologique* (2015) suggests *mettre à niveau* as a French equivalent for *upgrade* (*users can UPGRADE their operating system only to the Professional edition*). However, *upgrade* has three arguments (***X** upgrades **Y** to **Z***) while *mettre à niveau* has two arguments (**X** met **Y** à niveau).
- The syntactic constructions in which arguments are realized can also differ from one language to another. For example, the English adjective *compatible* can be used in three different syntactic structures: ***X** is compatible with **Y**; **Y**-compatible **X**; compatible **X***). In French, the equivalent *compatible* allows only two structures (***X** est compatible avec **Y**, **X** compatible*). Another example is the way some arguments can be permuted in a language; but not in another. For instance, two of the arguments of *search* in computing can be permuted easily in English (*a user SEARCHES information on the Internet; a user SEARCHES the Internet for information*); this permutation is not allowed in French (*un utilisateur RECHERCHE de l'information dans Internet; *un utilisateur RECHERCHE Internet pour de l'information*).
- Finally, some alternations (see Section 5.5) might exist in language A but not in language B. In English, *load* accepts an inchoative/causative alternation (*the program LOADS; the user LOADS a program*). In French, the verb *charger* can only be used in the causative sense (*CHARGER un programme; ?le programme SE CHARGE; ?le programme CHARGE*).

9.4 Equivalence in running text

Another way to consider ***equivalence relations*** between terms consists in examining how terms that appear in texts written one language are translated in texts in another language. Some researchers investigated this phenomenon, especially those interested in developing methods to extract terminological equivalents automatically.

A first problem is linked to structural differences between equivalents.[95] This was shown in a study conducted by Gaussier (2001) who looked at noun phrases

95. This problem differs from the structural differences examined in Section 9.3.3, since only nouns and noun phrases are considered here.

with two content words in English and French and characterized their correspondence based on morpho-syntactic patterns. His goal was to show how different structures in two languages impact automatic extraction. The author made a list of corresponding and non-corresponding patterns. Table 9.1 shows that English N + N terms can be rendered by a variety of different structures in French: N de N, N Prep N, N Adj, NN or N. This applies to all the other structures studied. Furthermore, Gaussier only considered terms with two content words. Matters can quickly become more complicated if longer multi-word terms are taken into consideration.

Table 9.1 Patterns of correspondences and non-correspondences (based on Gaussier 2001: 173, cited by Le Serrec et al. 2010: 82)

		ENGLISH				
		N N	**Adj N**	**N of N**	**N's N**	**N**
FRENCH	N de N	122	15	2	1	8
	N prep N	28	9	–	–	2
	N Adj	23	63	–	–	1
	N N	11	–	–	–	–
	N	1	1	–	–	–

When analyzing bilingual texts and the terms they contain, other divergences are likely to appear. First, a term in language A can have more than one equivalent in language B. Conversely, a term in language B can also be rendered by more than one term in language A. This problem is caused by ***variation*** in different languages. Carreño (2004) studied variation in an English/Spanish parallel corpus and showed that it affects almost all the studied terms, and that variants occur in different forms. Since variation occurs in all languages, it inevitably affects the identification of equivalents.

It is customary to consider that terms are equivalents if they belong to the same part of speech. These are the equivalents that terminologists record in terminological resources. However, in running text, terms that belong to given parts of speech may be rendered by terms that belong to different ones. For example, the French adjective term *climatique* will almost always be translated by the noun *climate* in English (Le Serrec 2008) even if the adjective *climatic* exists. Similarly, some meanings expressed by verbs in English may be expressed in French with nominalizations: *to execute a program* is often rendered in French by *l'exécution d'un programme*.

Finally, terms expressed in language A may be translated by an anaphora (a more generic term or a pronoun) in language B. For example, *the disk drive* in *the*

*disk drive is identified…*can be translated by *…ce dispositif est identifié…* instead of *l'unité de disque* if *unité de disque* was mentioned previously in the French text.

Summary

Equivalence is a relation established between terms that belong to different languages. When considered from a knowledge-based approach, the focus is placed on finding designations that label the same concept. When considered from the a lexical perspective, the meanings of terms guide the establishment of a possible equivalence.

In both perspectives, exact equivalence is sought since it is this form of equivalence recorded in terminological resources (dictionaries, term banks, thesauri, etc.). However, different situations can prevent terminologists from finding proper equivalents: non-equivalence, partial equivalence (or animorphism), structural differences between languages, etc.

Finally, equivalence can be established between terms in running texts. Other difficulties arise when attempting to do so, especially with automated methods. These difficulties are structural mismatches between equivalents in different languages, variation in both languages, non-correspondence between parts of speech and anaphora.

Further reading

For a broad overview of equivalence, see Adamska-Salaciak (2010). For a review on different approaches to equivalence in terminology, conceptual, lexical, and corpus-based, read Le Serrec et al. (2010).

Arntz (1993), Rondeau (1981) and Sager (1994) examine equivalence from a knowledge-based perspective. Fontenelle (2014, 2016) offers interesting accounts of having to deal with several languages in a term bank. Snow (1989, 2010) reports on the matter of adapting Common Law terminology to French. Office québécois de la langue française (2017) helps understand how an official organization plans the integration or replacement of terms borrowed from other languages.

Some authors explore the internal structure of multiword units in specialized domains and different languages (Rosario and Hearst 2001; Bouillon et al. 2012; Cabezas-García and Faber 2017).

Van Campenhoudt (1996) is one of the first authors who explored terminological equivalence from a lexicon-based perspective and reviewed some problems when establishing correspondences between terms.

Gaussier (2001) presents a comparison of structural differences between noun terms and describes challenges raised by the automated idenstification of terminological equivalents.

References

Adamska-Salaciak, A. 2010. "Examining Equivalence." *International Journal of Lexicography* 23(4): 387–409. https://doi.org/10.1093/ijl/ecq024

Aguado de Cea, G., and E. Montiel-Ponsoda. 2012. "Term variants in Ontologies." In *XXX Congreso Internacional de AESLA 2012*, 436–443. Universidad de Lleida, Spain.

Ahmad, K., and M. Rogers. 2001. "Corpus Linguistics and Terminology Extraction." In *Handbook of Terminology Management*, ed. by S. E. Wright, and G. Budin, 725–760. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.htm2.28ahm

Aitchison, J., D. Bawden, and A. Gilchrist. 2000. *Thesaurus Construction and Use: A Practical Manual*, 3rd edition. Chicago/London: Fitzroy Dearborn.

Aldestein, A., and M. T. Cabré. 2002. "The Specificy of Units with Specialized Meaning: Polysemy as an Explanatory Factor." *D.E.L.T.A.* 18: 1–25.

Alonso Campo, A. 2008. "Environmental Terminology in General Dictionaries." In *Proceedings of the XIII Euralex International Congress*, ed. by E. Bernal, and J. DeCesaris, 929–935. Barcelona: Institut universitari de lingüística aplicada (IULA).

Alonso Ramos, M. 2004. "Elaboración del *Diccionario de colocaciones del español y sus aplicaciones*." In *De Lexicografía. Actes del I Simposium Internacional de Lexicografía*, ed. by P. Bataner, and J. de Cesaris, 149–162. Barcelona: IULA & Edicions Petició.

Amsili, P. 2003. "L'antonymie en terminologie: quelques remarques." In *Terminologie et Intelligence Artificielle (TIA)*, 31–40. Strasbourg, France.

Arntz, R. 1993. "Terminological Equivalence and Translation." In *Terminology: Applications in Interdisciplinary Communication*, 5–20. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.70.02arn

Apresjan, J. 1973. "Regular Polysemy." *Linguistics* 142: 5–32.

BabelNet (http://babelnet.org/). Accessed 19 October 2017.

Barrière, C. 2002. "Investigating the Causal Relation in Informative Texts." *Terminology* 7(2): 135–154. https://doi.org/10.1075/term.7.2.02bar

Barrière, C., and A. Agbago. 2006. "TerminoWeb: A Software Environment for Term Study in Rich Contexts." In *Proceedings of the International Conference on Terminology, Standardisation and Technology Transfer* (TSTT 2006, Beijing), National Research Council of Canada.

Béjoint, H., and P. Thoiron (eds.). 2000. *Le sens en terminologie*. Lyon: Presses universitaires de Lyon.

Bergenholtz, H., and U. Kaufmann. (1997). "Terminography and Lexicography. A Critical Survey of Dictionaries from a Single Specialised Field." *Hermes* 18: 91–126.

Bernier-Colborne, G. 2016. *Aide à l'identification de relations lexicales au moyen de la sémantique distributionnelle et son application à un corpus bilingue du domaine de l'environnement*. PhD thesis presented at the University of Montreal, Montreal.

Binon, J., S. Verlinde, J. Van Dyck, and A. Bertels. 2000. *Dictionnaire d'apprentissage du français des affaires*. Paris: Didier.

Boas, H. C. 2009. "Semantic Frames as Interlingual Representations for Multilingual Lexical Databases." In *Multilingual FrameNets in Computational Lexicography: Methods and Applications*, ed. by H. C. Boas, 59–99. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110212976

Bouillon, P., E. Jezek, C. Melloni, and A. Picton, 2012. "Annotating Qualia Relations in Italian and French Complex Nominals." In *Language Resources and Evaluation Conference* (LREC 2012), 1527–1532. Istanbul, Turkey.

Boulanger, J. C. 1995. "Présentation : images et parcours de la socioterminologie." *Meta* 40(2): 194–205. https://doi.org/10.7202/002117ar

Bourigault, D., and M. Slodzian. 1999. "Pour une terminologie textuelle." *Terminologies nouvelles* 19: 29–32.

Bowker, L. 1993. "Multidimensional Classification of Concepts for Terminological Purposes." In *4th ASIS SIG/CR Classification Research Workshop*, 39–56. Columbus, Ohio.

Bowker, L. 1996. "Learning from Cognitive Science: Developing a New Approach to Classification in Terminology." In *Euralex '96 Proceedings*, ed. by M. Gellerstam, J. Järborg, S. G. Malmgren, K. Norén, L. Rogström, and C. Röjder, 781–788. Gothenburg, Sweden.

Bowker, L., and J. Pearson. 2002. *Working with Specialized Languages. A Practical Guide to Using Corpora*. London: Routledge. https://doi.org/10.4324/9780203469255

Cabezas-García, M., and P. Faber. 2017. "A Semantic Approach to the Inclusion of Complex Nominals in English Terminographic Resources." In *Computational and Corpus-Based Phraseology. EUROPHRAS 2017*. Lecture Notes in Computer Science, ed. by R. Mitkov, 145–159. Cham: Springer. https://doi.org/10.1007/978-3-319-69805-2_11

Cabré, M. T. 1999. *Terminology. Theory, Methods and Applications*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/tlrp.1

Cabré, M. T. 2003. "Theories of Terminology: Their Description, Prescription and Explanation." *Terminology* 9(2): 163–200. https://doi.org/10.1075/term.9.2.03cab

Campo, A. 2013. *The Reception of Eugen Wüster's Work and the Development of Terminology*. PhD thesis presented at the University of Montreal, Montreal.

Candito M., P. Amsili, L. Barque, F. Benamara, G. Chalendar, M. Djemaa, P. Haas, R. Huyghe, Y. Mathieu, P. Muller, B. Sagot, and L. Vieu. 2014. "Developing a French FrameNet: Methodology and First Results." In *Proceedings of the Ninth International Conference on Language Resources and Evaluation* (LREC), Reykjavik, Iceland (http://www.lrec-conf.org/proceedings/lrec2014/index.html). Accessed 18 April 2019.

Carreño, I. 2004. *Analyse de la variation terminologique en corpus parallèle anglais-espagnol et de son incidence sur l'extraction de termes bilingues*. Master's thesis presented at the University of Montreal, Montreal.

Carrière, I. 2006. *Adjectifs dérivés de noms : analyse en corpus médical et modèle d'encodage terminologique*. Master's thesis presented at the University of Montreal, Montreal

Cohen, B. 1986. *Lexique de cooccurrents: bourse et conjoncture économique*. Montréal: Linguatech.

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.

Condamines, A. 2005. "The Interpretation of Corpus Semantics: The Case of Terminology Construction." *Revue de linguistique appliquée* 12(1): 39–52.

Coxhead, A. 2000. "A New Academic Word List." *TESOL Quarterly* 34(2): 213–238. https://doi.org/10.2307/3587951

Croft, W., and D. A. Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803864

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

Cruse, A. 2011. *Meaning in Language: An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.

Daille, B. 2001. "Qualitative Term Extraction." In *Recent Advances in Computational Terminology*, ed. by D. Bourigault, C. Jacquemin, and M. C. L'Homme, 149–166. Amsterdam/Philadephia: John Benjamins. https://doi.org/10.1075/nlp.2.08dai

Daille, B. 2017. *Term Variation in Specialised Corpora: Characterisation, Automatic Discovery and Applications*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/tlrp.19

Daille, B., B. Habert, C. Jacquemin, and J. Royauté. 1996. "Empirical Observation of Term Variation and Principles for their Description." *Terminology* 3(2): 197–257. https://doi.org/10.1075/term.3.2.02dai

Dannélls, D., K. Friberg Heppin, and A. Ehrlemark. 2014. "Using Language Technology Resources and Tools to Construct Swedish FrameNet." In *Proceedings of the Workshop on Lexical and Grammatical Resources for Language Processing*, 8–17. Dublin, Ireland. https://doi.org/10.3115/v1/W14-5802

De Saussure, F. 1916. *Cours de linguistique générale*. Paris: Payot.

DICE. *Diccionario de Colocaciones del Español* (http://www.dicesp.com/paginas). Accessed 4 April 2018.

Dicionário da Copa do Mundo (https://www.ufjf.br/framenetbr/dicionario/). Accessed 10 October 2019.

DiCoEnviro. 2018. *Dictionnaire fondamental de l'environnement* (http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search.cgi). Accessed 4 April 2018.

DicoInfo. 2018. *Dictionnaire fondamental de l'informatique et de l'Internet* (http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi). Accessed 4 April 2018.

Dicouèbe. Le DiCo en ligne (http://olst.ling.umontreal.ca/dicouebe/main.php). Accessed 4 April 2018.

Diki-Kidiri, M. 2000. "Une approche culturelle de la terminologie." *Terminologies nouvelles* 21: 27–41.

Diki-Kidiri, M. 2007. "Éléments de terminologie culturelle." *Les cahiers du RIFAL* 26: 14–25.

Djemma, M. 2017. *Stratégie domaine par domaine pour la création d'un FrameNet du français : annotations en corpus de cadres et rôles sémantiques*. PhD thesis presented at Université Paris-Diderot, Paris.

Dolbey, A., M. Ellsworh, and J. Scheffczyk. 2006. "BioFrameNet: A Domain-specific FrameNet Extension with Links to Biomedical Ontologies." In *KR-MED 2006*, *Biomedical Ontology in Action*, 87–94. Baltimore, Maryland.

Drouin, P. 2007. "Identification automatique du lexique scientifique transdisciplinaire." *Revue française de linguistique appliquée* 12(2): 45–64.

Drouin, P., A. Francoeur, J. Humbley, and A. Picton (eds.). 2017. *Currents Trends in Terminological Variation*. Amsterdam/Philadephia: John Benjamins.

Dubuc, R. 2002. *Manuel pratique de terminologie*. Brossard (Québec): Linguatech.

Durán-Muñoz, I., and M. Bautista-Zambrana. 2017. "Applying Ontologies to Terminology: Advantages and Disadvantages." *HERMES – Journal of Language and Communication in Business* 51: 65–77. https://doi.org/10.7146/hjlcb.v26i51.97438

Dury, P. (1999). "Étude comparative et diachronique des concepts ecosystem et écosystème." *Meta* 44(3): 485–499. https://doi.org/10.7202/002690ar

EcoLexicon. 2018. *Terminological Knowledge Base* (http://ecolexicon.ugr.es/en/index.htm). Accessed 18 July 2018.

ENVO. Environmental Ontology. 2015. (http://environmentontology.org/). Accessed 19 February 2015.

Estopà, R. 2001. "Les unités de signification spécialisées élargissant l'objet du travail en terminologie." *Terminology* 7(2): 217–237. https://doi.org/10.1075/term.7.2.06est

Faber, P. (ed.). 2012. *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110277203

Faber, P. 2014. "Frames as a Framework for Terminology." In *Handbook in Terminology*, ed. by H. J. Kockert, and F. Steurs, 14–33. Amsterdam/Philadelphia: John Benjamins.

Faber, P., P. León-Araúz, and A. Reimerink. 2016. "EcoLexicon: New features and Challenges." In *GLOBALEX 2016: Lexicographic Resources for Human Language Technology* in conjunction with the 10th edition of the *Language Resources and Evaluation Conference*, ed. by I. Kernerman, I. Kosem Trojina, S. Krek, and L. Trap-Jensen, 73–80. Portorož, Slovenia.

Faber, P., J. Verdejo, P. León Araúz, A. Reimerink, and G. Guzmán. 2014. "Neural Substrates of Specialized Knowledge Representation: An fMRI study." *Revue française de linguistique appliquée* 19(1): 15–32.

Felber, H. 1984. *Terminology Manual*. Paris: Unesco and Infoterm.

Fellbaum, C. (ed.). 1998. *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/7287.001.0001

Fillmore, C. J. 1968. "The Case for Case." In *Universals in Linguistic Theory*, ed. by E. Bach, and R. T. Harns, 1–88. New York: Holt, Rinehard & Winston.

Fillmore, C. J. 1976. "Frame Semantics and the Nature of Language." *Annals New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280: 20–32. https://doi.org/10.1111/j.1749-6632.1976.tb25467.x

Fillmore, C. J. 1982. "Frame Semantics." In *Linguistics in the Morning Calm*, ed. by The Linguistic Society of Korea, 111–137. Seoul: Hanshin.

Fillmore, C. J. 1985. "Frames and the Semantics of Understanding." *Quaderni di Semantica* 6: 222–254.

Fillmore, C. J., and B. T. Atkins. 1992. "Toward a Frame-based Lexicon: The Semantics of RISK and its Neighbors." In *Frames, Fields and Contrasts*, ed. by A. Lehrer, and E. Feder Kittay, 75–102. Hillsdale, New Jersey: Lawrence Erlbaum.

Fillmore, C. J., and C. Baker. 2010. "A Frames Approach to Semantic Analysis." In *The Oxford Handbook of Linguistic Analysis*, ed. by B. Heine, and H. Narrog, 313-339. Oxford: Oxford University Press.

Fillmore, C. J., C. R. Johnson, and M. Petruck. 2003. "Background to FrameNet." *International Journal of Lexicography* 16(3): 235–250. https://doi.org/10.1093/ijl/16.3.235

FOLDOC, A Free Online Dictionary of Computing (foldoc.org). Accessed 17 February 2015.

Fontenelle, T. 1997. *Turning a Bilingual Dictionary into a Lexico-semantic Database*. Tübingen: Niemeyer. https://doi.org/10.1515/9783110920116

Fontenelle, T. (ed.). 2003. *FrameNet and Frame Semantics. Special Issue of the International Journal of Lexicography* 16(3).

Fontenelle, T. 2014. "From Lexicography to Terminology: A Cline, not a Dichotomy." In *Euralex 2014. Proceedings*, 25–46. Bolzano, Italy.

Fontenelle, T. 2016. "Traduction et interprétation face aux défis actuels." *Revue française de linguistique appliquée* 21(1): 53–66.

Foy, N. F., and D. L. McGuiness. 2014. *Ontology Development 101: A Guide to Creating Your First Ontology* (https://protegewiki.stanford.edu/wiki/Ontology101). Accessed 4 July 2018.

A Framed Version of DiCoEnviro (http://olst.ling.umontreal.ca/dicoenviro/framed/index.php). Accessed 4 April 2018.

FrameNet (https://framenet.icsi.berkeley.edu/fndrupal/). Accessed 10 October 2019.

Frawley, W. 1988. "New Forms of Specialized Dictionaries." *International Journal of Lexicography* 1(3): 189–213. https://doi.org/10.1093/ijl/1.3.189

Freixa, J. 2006. "Causes of Denominative Variation in Terminology." *Terminology* 12(1): 51–77. https://doi.org/10.1075/term.12.1.04fre

Gagné, A. M., and M. C. L' Homme. 2016. "Opposite Relationships in Terminology." *Terminology* 22(1): 30–51. https://doi.org/10.1075/term.22.1.02gag

Gaudin, F. 1993. *Pour une socioterminologie: des problèmes sémantiques aux pratiques institutionnelles*. Rouen: Presses de l'Université de Rouen.

Gaudin, F. 2003. *Socioterminologie: une approche sociolinguistique de la terminologie*. Bruxelles: Duculot. https://doi.org/10.3917/dbu.gaudi.2003.01

Gaussier, É. 2001. "General Considerations on Bilingual Terminology Extraction." In *Recent Advances in Computational Terminology*, ed. by D. Bourigault, C. Jacquemin, and M. C. L' Homme, 167–182. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/nlp.2.09gau

Geerearts, D. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.

GEMET Thesaurus. (http://www.eionet.europa.eu/gemet/). Accessed 19 February 2015.

Gentilhomme, Y. 1994. "L'éclatement du signifié dans les discours technoscientifiques." *Cahiers de lexicologie* 54(1): 5–35.

Ghazzawi, N. 2016. *Du terme prédicatif au cadre sémantique : méthodologie de compilation d'une ressource terminologique pour des termes prédicatifs arabes en informatique*. PhD thesis presented at the University of Montreal, Montreal.

Le grand dictionnaire terminologique (http://www.granddictionnaire.com/). Accessed 26 February 2018.

Grass, T. 1999. *La traduction juridique bilingue français-allemand: problématique et résolutions des ambiguïtés terminologiques*. Bonn: Romanistischer Verlag.

Gross, G. 1994. "Classes d'objets et descriptions des verbes." *Langages* 115: 15–30. https://doi.org/10.3406/lgge.1994.1684

Guilbert, L. 1973. "La spécificité du terme scientifique et technique." *Langue française* 17: 5–17. https://doi.org/10.3406/lfr.1973.5617

Granger, S., and M. Paquot. 2009. "In Search of a General Academic English: A Corpus-driven Study." In *Proceedings of the International Conference on LSP. Options and Practices of LSP Practitioners*, Heraklion.

Hanks, P., and J. Pustejovksy. 2005. "A Pattern Dictionary for Natural Language Processing." *Revue française de linguistique appliquée* 10(2): 63–82.

Harris, Z. 1968. *Mathematical Structures of Language*. New York: Wiley.

Hatier, S. 2016. *Identification et analyse linguistique du lexique scientifique transdisciplinaire. Approche outillée sur un corpus d'articles de recherche en SHS*. PhD thesis presented at Université Grenoble-Alpes, Grenoble.

Hirst, G. 2009. "Ontology and the Lexicon." In *Handbook on Ontologies*, ed. by S. Staab, and R. Studer, 269–292, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-92673-3_12

Humbley, J. 2007. "Vers une réception plurielle de la théorie terminologique de Wüster: une lecture commentée des avant-propos successifs du manuel Einfürhung in die allgemeine Terminologie Lehre." *Langages* 168(4): 82–91. https://doi.org/10.3917/lang.168.0082

Ingrosso, F., and A. Polguère. 2015. "How Terms Meet in Small-World Lexical Networks: The Case of Chemistry Terminology." In *Proceedings of the Conference Terminology and Artificial Intelligence 2015*, 167–171, Granada, Spain.

International Commission on Zoological Nomenclature. 2000. *International Code of Zoological Nomenclature*, 4th Edition (http://www.nhm.ac.uk/hosted-sites/iczn/code/). Accessed 28 August 2012.

Josselin-Leray, A. 2005. *Place et rôle des terminologies dans les dictionnaires unilingues et bilingues. Étude d'un domaine de spécialité: vulcanologie*. PhD thesis presented at the University Lumière Lyon-2, Lyon.

Jousse, A. L., and A. Polguère. 2005. *Le DiCo et sa version Dicouèbe. Document descriptif et manuel d'utilisation*. Université de Montréal : Département de linguistique et de traduction.

Kageura, K. 2002. *The Dynamics of Terminology: A Descriptive Theory of Term Formation and Terminological Growth*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/tlrp.5

Kageura, K. 2012. *The Quantitative Analysis of the Dynamics and Structure of Terminologies*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/tlrp.15

Kicktionary. *The Multilingual Electronic Dictionary of Football Language* (http://www.kicktionary.de/index.html). Accessed 16 May 2017.

Kleiber, G. 1999. *Problèmes de sémantique. La polysémie en question*. Villeneuve d'Ascq : Presses universitaires du Septentrion.

Kocourek, R. 1991. *La langue française de la technique et de la science. Vers une linguistique de la langue savante*. Niemeyer: Oscar Brandstetter.

L'Homme, M. C. 2002. "Fonctions lexicales pour représenter les relations sémantiques entre termes." *Traitement automatique de la langue (TAL)* 43(1): 19–41.

L'Homme, M. C. 2003. "Capturing the Lexical Structure in Special Subject Fields with Verbs and Verbal Derivatives: A Model for Specialized Lexicography." *International Journal of Lexicography* 16(4): 403–422. https://doi.org/10.1093/ijl/16.4.403

L'Homme, M. C. 2004a. "A Lexico-semantic Approach to the Structuring of Terminology." In *Computerm 2004. Workshop help in conjunction with Coling 2004*, 7–14. University of Geneva, Geneva, Switzerland.

L'Homme, M. C. 2004b. *La terminologie: principes et techniques*. Montréal: Les Presses de l'Université de Montréal. https://doi.org/10.4000/books.pum.10693

L'Homme, M.C. 2005. "Sur la notion de 'terme.'" *Meta* 50(4): 1112-1132.

L'Homme, M. C. 2007. "Using Explanatory and Combinatorial Lexicology to Describe Terms." In *Selected Lexical and Grammatical Topics in the Meaning-Text Theory*. In Honour of Igor Mel'cuk, ed. by L. Wanner, 163–198. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/slcs.84.06hom

L'Homme, M. C. 2012. "Using ECL (Explanatory Combinatorial Lexicology) to Discover the Lexical Structure of Specialized Subject Fields." In *Words, Meanings and other Interesting Things. A Festschrift n Honour of the 80th Anniversary of Professor Igor Alexandrovic Mel'cuk*, ed. by J. Apresjan, et al., 378–390. Moscow: RCK.

L'Homme, M. C. 2015a. "Terminologies and Taxonomies." In *The Handbook of the Word*, ed. by J. Taylor, 334–349. Oxford: Oxford University Press.

L'Homme, M. C. 2015b. "Predicative Lexical Units in Terminology." In *Recent Advances in Language Production, Cognition and the Lexicon*, ed. by N. Gala, R. Rapp, and G. Bel-Enguix, 75–93. Berlin: Springer.

L'Homme, M. C. 2018. "Maintaining the Balance between Knowledge and the Lexicon in Terminology: A Methodology based on Frame Semantics." In *Medical Lexicography and Terminology. Special issue of Lexicography*, ed. by P. Peters, J. G. Yongwei, and J. Ding, 4(1): 3–21.

L'Homme, M. C., and G. Bernier-Colborne. 2012. "Terms as Labels for Concepts, Terms as Lexical Units: A Comparative Analysis in Ontologies and Specialized Dictionaries." *Applied Ontology* 7(4): 387–400.

L'Homme, M. C., and A. Polguère. 2008. "Mettre en bons termes les dictionnaires spécialisés et les dictionnaires de langue générale." In *Lexicographie et terminologie : histoire de mots. Hommage à Henri Béjoint*, ed. by F. Maniez, and P. Dury, 191–206. Lyon: Presses de l'Université de Lyon.

L'Homme, M. C., B. Robichaud, and N. Prévil. 2018. "Browsing the Terminological Structure of a Specialized Domain: A Method Based on Lexical Functions and their Classification." *Language Resources and Evaluation, LREC 2018*, Myazaki, Japan (http://www.lrec-conf.org/proceedings/lrec2018/summaries/840.html). Accessed 18 April 2019.

L'Homme, M. C., and A. San Martín. 2016. "Définition terminologique : systématisation de règles de rédaction dans les domaines de l'informatique et de l'environnement." *Cahiers de lexicologie* 109: 147–174.

Le Serrec, A. 2008. *Étude sur l'équivalence de termes extraits automatiquement d'un corpus parallèle : contribution à l'extraction terminologique bilingue*. Master's thesis presented at the University of Montreal, Montreal.

Le Serrec, A., M. C. L'Homme, P. Drouin, and O. Kraif. 2010. "Automating the Compilation of Specialized Dictionaries: Use and Analysis of Term Extraction and Lexical Alignment." *Terminology* 16(1): 77–106. https://doi.org/10.1075/term.16.1.04les

Legal Dictionary, Law.com (dictionary.law.com). Accessed 29 July 2015.

Lehrberger, J. 2003. "Automatic Translation and the Concept of Sublanguage." In *Readings in Machine Translation*, ed. by S. Nirenburg, H. Somers, and Y. Wilks, 207–220. Boston: Massachussetts Institute of Technology.

León Araúz, P. 2009. *Representación multidimensional del conocimiento especializado: el uso de marcos desde la macroestructura hasta la microestructura*. PhD thesis presented at the University of Granada, Granada.

León Araúz, P., P. Faber, and S. Montero Martínez. 2012. "Specialized Language Semantics." In *A Cognitive Linguistics View of Terminology and Specialized Language*, ed. by P. Faber, 97–175. Berlin: De Gruyter.

León Araúz, P., and A. Reimerink. 2010. "Knowledge Extraction and Multidimensionality in the Environmental Domain." In *Proceedings of the Terminology and Knowledge Engineering (TKE) Conference*. Dublin: Dublin City University (http://lexicon.ugr.es/pdf/leonreimerink2010.pdf). Accessed 18 April 2019.

Lerat, P. 2002a. "Vocabulaire juridique et *schémas* d'arguments juridiques." *Meta* 47(2): 155–162. https://doi.org/10.7202/008005ar

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

Longman Dictionaries. 2008. *Longman American Defining Vocabulary*. (http://www.webcitation.org/6O48nGWbq). Accessed 14 February 2014.

Lorente, M. 2007. "Les unitats lèxiques verbals dels textos especialitzats. Redefinició d'una proposta de classificació." In *Estudis de lingüístics i de lingüística aplicada en honor de M. Teresa Cabré Catellví. Volum II: De deixebles* (Sèrie Monografies 11–12), ed. by M. Lorente, et al., 365–380. Barcelona: Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra.

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

Malm, P., S. Mumtaz Virk, L. Borin, and A. Savera. 2018. "LingFN. Towards a FrameNet for the Linguistics Domain." In *International FrameNet Workshop. 2018. Multilingual FrameNets and Constructions. Proceedings. Language Resources and Evaluation* (LREC 2018), ed. by T. Timponi Torrent, L. Borin, and C. Baker, 37–43. Miyazaki, Japan.

Maniez, F. 2015. "The Use of Denominal Adjectives in Terminology." In *Terminología y comunicación científica y social*, ed. by M. Roldán Vendrell, 43–66. Granada: Comares.

Marcel, E. 2000. "Quelques modalités d'application de la fonction lexicale Gener en botanique." *Cahiers de lexicologie* 76(1): 91–110.

Marshman E. 2006. *Lexical Knowledge Patterns for Semi-automatic Extraction of Cause-effect and Association Relations from Medical Texts: A Comparative Study of English and French*. PhD thesis presented at the University of Montreal, Montreal.

Mel'čuk, I. A. 2004. "Actants in Semantics and Syntax I: Actants in Semantics." *Linguistics* 42(1): 1–66. https://doi.org/10.1515/ling.2004.004

Mel'čuk, I. A. 2013. *Semantics. From Meaning to Text*, Vol 2. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/slcs.135

Mel'čuk. I. et al. 1984–1999. *Dictionnaire explicatif et combinatoire*. Montréal : Presses de l'Université de Montréal.

Mel'čuk. I., A. Clas, and A. Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Duculot: Louvain-la-Neuve.

Mel'čuk, I., and J. Milicevic. 2014. *Introduction à la linguistique*, Vol. 1. Paris: Hermann.

Mel'čuk, I., and A. Polguère. 2007. *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Bruxelles: De Boeck.

Merriam-Webster Online Dictionary (https://www.merriam-webster.com/). Accessed 4 April 2018.

Meyer, I., D. Skuce, L. Bowker, and K. Eck. 1992. "Towards a New Generation of Terminological Resources: An Experiment in Building a Terminological Knowledge Base." In *Proceedings of the 16th International Conference on Computational Linguistics*, 956–957, Nantes, France. https://doi.org/10.3115/992383.992410

Meyer, I., and K. Mackintosh. 1996. "The Corpus from a Terminographer's Viewpoint." *International Journal of Corpus Linguistics* 1(2): 257–268. https://doi.org/10.1075/ijcl.1.2.05mey

Meyer, I., and K. Mackintosh. 2000. "When Terms Move into our Everyday Lives: An Overview of De-terminologization." *Terminology* 6(1): 111–138. https://doi.org/10.1075/term.6.1.07mey

Murphy, M. L. 2003. *Semantic Relations in the Lexicon: Antonymy, Synonymy and other Paradigms*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486494

Navigli, R., and S. Ponzetto. 2012. "BabelNet: The Automatic Construction, Evaluation and Application of a Wide-coverage Multilingual Semantic Network." *Artificial Intelligence* 193: 217–250. https://doi.org/10.1016/j.artint.2012.07.001

NeoVisual (http://olst.ling.umontreal.ca/dicoenviro/neovisuel/index.php). Accessed 19 October 2018.

Nuopponen, A. 1994. "Wüster Revisited: On Causal Concept Relationships and Causal Concept Systems." In *Applications and Implications of Current LSP Research, Proceedings of the 9th European Symposium on LSP*, Vol. II, 532–539. Bergen: Fagbokforlaget.

Nuopponen, A. 2014. "Tangled Web of Concept Relations. Concept relations for ISO 1087-1 and ISO 704." In *Terminology and Knowledge Engineering* 2014. Berlin, Germany.

Office québécois de la langue française. 2015. *À vos vélos* (https://www.oqlf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/terminologie_velo/fiches/indexFRA.html). Accessed 27 July 2015.

Office québécois de la langue française. 2017. *Politique de l'emprunt linguistique* (https://www.oqlf.gouv.qc.ca/ressources/bibliotheque/terminologie/20170330_politique_emprunt.pdf). Accessed 25 February 2017.

Ohara, K. 2009. "Frame-based Contrastive Lexical Semantics in Japanese FrameNet: The Case of *Risk* and *Kakeru*." In *Multilingual FrameNets in Computational Lexicography,* ed. by H. C. Boas, 163–182. New York/Berlin: Mouton De Gruyter.

Palmer, F. R. 1976. *Semantics. A New Outline*. Cambridge: Cambridge University Press.

Pavel, S., and D. Nolet. 2001. *Handbook of Terminology*. Ottawa: Public Works and Government Services Canada. Translation Bureau.

Pecman, M. 2018. *Langue et construction de connaisSENSes : énergie lexico-discursive et potentiel sémiotique des sciences*. Paris: L'Harmattan.

Pedersen, B. S., S. Nimb, A. Søgaard, M. Hartmann, and S. Olsen. 2018. "A Danish FrameNet Lexicon and an Annotated Corpus Used for Training and Evaluating a Semantic Frame Classifier." In *Language Resources and Evaluation, LREC 2018*, Myazaki, Japan (http://www.lrec-conf.org/proceedings/lrec2018/summaries/586.html). Accessed 18 April 2019.

Petit, G. 2000. "L'introuvable identité du terme technique." *Revue française de linguistique appliquée* 6(2): 63–79.

Phal, A. 1971. *Vocabulaire général d'orientation scientifique*. Paris: CRÉDIF.

Picht, H., and J. Draskau. 1985. *Terminology: An Introduction*. Guilford: University of Surrey.

Pimentel, J. 2013. "Methodological Bases for Assigning Terminological Equivalents. A Contribution." *Terminology* 19(2): 237–257. https://doi.org/10.1075/term.19.2.04pim

Pitkänen-Heikkilä, P. 2015. "Adjectives as Terms." *Terminology* 21(1): 76–101. https://doi.org/10.1075/term.21.1.04pit

Polguère, A. 2012. "Propriétés sémantiques et combinatoires des quasi-prédicats sémantiques." *Scolia* 26: 131–152.

Polguère, A. 2014. "From Writing Dictionaries to Weaving Lexical Networks." *International Journal of Lexicography* 27(4): 396–418. https://doi.org/10.1093/ijl/ecu017

Polguère, A. 2016. *Lexicologie et sémantique lexicale : notions fondamentales*. Montréal: Presses de l'Université de Montréal. https://doi.org/10.2307/j.ctv69t90p

Polguère, A., and D. Sikora (eds.). 2016. La definition, special issue of *Cahiers de lexicologie* 109(2).

PropBank. The Proposition Bank (http://propbank.github.io/). Accessed 20 July 2018.

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

Riemer, N. 2015. Word Meanings." In *The Handbook of the Word*, ed. by J. Taylor, 305–319. Oxford: Oxford University Press.

Rey, A. 1976. *Terminologie : noms et notions*, coll. Que sais-je. Paris: Presses universitaires de France.

Robichaud, B. 2012. "Logic Based Methods for Terminological Assessment." In *Proceedings of the Eight International Conference on Language Resources and Evaluation* (LREC'12), ed. by N. Calzolari et al., 94–98. Istanbul, Turkey.

Roche, C. 2012. "Ontoterminology: How to Unify Terminology and Ontology into a Single Paradigm." In *Language Resources and Evaluation (LREC 2012). Proceedings*, 2626–2630. Istanbul, Turkey.

Rondeau, G. 1981. *Introduction à la terminologie*. Montréal: Centre éducatif et culturel.

Rondeau, G. 1984. *Introduction à la terminologie*. Chicoutimi: Gaëtan Morin éditeur.

Rosario, B., and M. Hearst. 2001. "Classifying the Semantic Relations in Noun Compounds via a Domain-specific Lexical Hierarchy." In *Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing (EMNLP-01)*, 82–90. Cornell: Cornell University.

Rosch, E. 1978. "Principles of Categorization." In *Cognition and Categorization*, ed. by E. Rosch, and P. Lloyd, 27–48. Hillsdale, N.J.: Erlbaum.

Ruppenhofer, J, M. Ellsworth, M. Petruck, C. Johnson, C. Baker, and J. Scheffczyk, 2016. *FrameNet II: Extended Theory and Practice* (https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=the_book). Accessed 27 January 2017.

Sager, J. C. 1990. *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.44

Sager, J. C. 1994. "Reflections on Terminological Translation Equivalents." *Lebende Sprachen* 39(2): 55–57. https://doi.org/10.1515/les.1994.39.2.55

Sager, J. C. 1998. "Search of a Foundation towards a Theory of the Term." *Terminology* 5(1): 41–57. https://doi.org/10.1075/term.5.1.05sag

Sager, J. C. 2000. *Essays on Definition*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/tlrp.4

San Martín, A. 2016. *La representación de la variación contextual mediante definiciones terminológicas flexible*, PhD thesis presented at the University of Granada, Granada.

San Martín, A., and M. C. L'Homme. 2014. "Definition Patterns for Predicative Terms in Specialized Dictionaries." In *Language Resources and Evaluation, LREC 2014*, Reykjavik, Iceland (http://www.lrec-conf.org/proceedings/lrec2014/index.html). Accessed 10 May 2018.

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

Seppälä, S., A. Ruttenberg, Y. Schreiber, and B. Smith. 2016. "Definitions in Ontologies." *Cahiers de lexicologie* 109(2), 173–205.

Snow, G. 1989. "La "refrancisation" de la common law: entreprise néologique." In *Biennale de la langue française (1989). Exprimer la modernité en français*, 387–389. Montréal : Office de la langue française.

Snow, G. 2010. "La common law de A à Z, une ressource de plus pour les langagiers." *L'Actualité langagière* 7(4): 7–10.

Staab, S., and R. Studer (eds.). 2009. *Handbook on Ontologies*, 2nd Edition. Heidelberg: Springer. https://doi.org/10.1007/978-3-540-92673-3

Structural Informatics Group. 2018. *Foundational Model of Anatomy* (http://si.washington.edu/projects/fma). Accessed 31 May 2018.

Studer, R., R. Benjamins, and D. Fensel. 1998. "Knowledge Engineering: Principles and Methods." *Data & Knowledge Engineering* 25(1–2): 161–198. https://doi.org/10.1016/S0169-023X(97)00056-6

Subirats Rüggeberg, C. 2009. "Spanish FrameNet: An On-line Lexical Resource for Spanish." In *Multilingual FrameNets in Computational Lexicography. Methods and Applications*, ed. by H. Boas, 135–162. New York/ Berlín: Mouton de Gruyter.

Taylor, J. (ed.). 2015. *The Handbook of the Word*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641604.001.0001

Temmerman, R. 2000. *Towards New Ways of Terminological Description: The Sociocognitive Approach*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/tlrp.3

TERMIUM Plus® (http://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-fra.html?lang=fra). Accessed 4 August 2017.

Torrent, T. T., and M. Ellsworth. 2013. "Behind the Labels: Criteria for Defining Analytical Categories in FrameNet Brasil." *Revista Veredas* 17(1): 44–65.

Tutin, A. 2007. "Autour du lexique et de la phraséologie des écrits scientifiques." *Revue française de linguistique appliquée* 12(2): 5–14.

Van Campenhoudt, M. 1996. "Réseau notionnel, intelligence artificielle et équivalence en terminologie multilingue : essai de modélisation." In *Lexicomatique et dictionnairiques, IVes journées scientifiques du réseau thématique Lexicologie, terminologie, traduction*, ed. by A. Clas, P. Thoiron, and H. Béjoint, 281–306. Montréal, AUPELF-UREF et Beyrouth: F.M.A.

Van Valin, R. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610578

Vendler, Z. 1967. *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.

VerbNet. A Class-based Verb Lexicon (https://verbs.colorado.edu/~mpalmer/projects/verbnet.html). Accessed 16 May 2017.

Vézina, R., X. Darras, J. Bédard, and M. Lapointe-Giguère. 2009. *La rédaction de définitions terminologiques*. Montréal: Office québécois de la langue française.

Wandji, O., M. C. L'Homme, and N. Grabar. 2013. "Discovery of Semantic Frames for a Contrastive Study of Verbs in Medical Corpora." *Terminology and Artificial Intelligence*, TIA 2013, Paris.

Wanner, L. 1996. *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/slcs.31

Webopedia (http://www.webopedia.com/). Accessed 29 July 2015.

Wierzbicka, A. 1985. *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma.

Winston, M. E., R. Chaffin, and D. Herrman. 1987. "A Taxonomy of Part-Whole Relationships." *Cognitive Science* 11(4): 417–444. https://doi.org/10.1207/s15516709cog1104_2

Wittgenstein, L. 1956[1972]. *Philosophical Investigations*, translated by G. G. M. Anscombre. Oxford: Blackwell.

WordNet. A Lexical Database for English (http://wordnet.princeton.edu/). Accessed 3 August 2015.

Wüster, E. 1968. *The Machine Tool Dictionary. An Interlingual Dictionary of Basics Concepts*. London: Technical Press.

Wüster, E. 1979. *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. Wien: Springer.

You, L., and K. Liu. 2005. "Building Chinese FrameNet." In *2005 International Conference on Natural Language Processing and Knowledge Engineering. Proceedings*, Wuhan, China (https://ieeexplore.ieee.org/document/1598752). Accessed 19 April 2019.

Corpus examples references

Axsen, J., C. Yang, R. W. McCarthy, A. F. Burke, K. S. Kurani, and T. S. Turrentine. 2011. "The Plug-in Electric Vehicle Pathway." In *Sustainable Transportation Energy Pathways: A Research Summary for Decision Makers* (file:///C:/Documents%20and%20Settings/p1005591/Mes%20documents/Downloads/2011-UCD-ITS-RP-11-67.pdf). Accessed 21 March 2014.

A Beginner's Guide to HTML. NCSA (http://archive.ncsa.uiuc.edu/General/Internet/WWW/HTMLPrimerAll.html). Accessed 23 Octobre 2001.

Benson, S., and R. Nagel. 2004. *Endangered Species*, 2nd Edition. Farmington Hills: Thomson Gale (http://www.about-the-web.com/). Accessed 15 August 2017.

Bjork, I., C. Connors, T. Welch, D. Shaw, and W. Hewitt. 2011. *Encouraging Renewable Energy Development. A Handbook for International Energy Regulators*, National Association of Regulatory Utility Commisioners (https://www.naruc.org/international/what-we-do/clean-energy/regulating-clean-energy-handbook/). Accessed 26 September 2019.

Brain, M. "How Computers Programs Work." *Howstuffworks* (http://www.howstuffworks.com/program.htm). Accessed 2 November 2001.

Brain, M. "How Microprocessors Work." *Howstuffworks* (http://www.howstuffworks.com/microprocessor.htm). Accessed 2 November 2001.

Brown, G. "How Sound Cards Work." *Howstuffworks* (http://www.howstuffworks.com/sound-card.htm). Accessed 1 November 2001.

Cabral, J. P. S. 2010. "Water Microbiology. Bacterial Pathogens ans Water." *International Journal of Environmental Research and Public Health* 7: 3657–3703 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2996186/). Accessed 12 May 2017. https://doi.org/10.3390/ijerph7103657

California Ocean Protection Council and California Ocean Science Trust. 2011. *Plastic Debris in the California Marine Ecosystem* (http://dornsife.usc.edu/assets/sites/291/docs/Plastic_Report_lowres.pdf). Accessed 12 May 2017.

Campbell, L. 1995. "Endangered and Threatened Animals of Texas: Their Life History and Management." *Texas Parks and Wildlife* (http://agris.fao.org/agris-search/search.do?recordID=US9631480). Accessed 28 February 2017.

Castro, A. *Yet Another Computer Users Guide* (http://www.geocities.com/compguide111/index.htm). Accessed 22 December 2001.

Cohen, L. *A Basic Guide to the Internet* (http://library.albany.edu/internet/internet.html). Accessed 5 August 2004.

Comer, D. E. *Questions and Answers on Networking and Internets* (http://www.netbook.cs.purdue.edu/othrpags/qanda.htm). Accessed 16 August 2004.

Computers for Beginners (http://tmbzone.com/index.htm). Accessed 4 August 2004.

Cooke, S. J. 2008. "Biotelemetry and Biologging in Endangered Species Research and Animal Conservation: Relevance to Regional, National and IUCN Red List Threat Assessments." *Endangered Species Research* 4: 165–185 (https://www.fecpl.ca/wp-content/uploads/2013/06/Cooke_ESR_Review.pdf). Accessed 27 June 2019. https://doi.org/10.3354/esr00063

Curtin, M. *Introduction to Network Security* (http%3A%2F%2Fwww.interhack.net%2Fpubs%2Fnetwork-security%2F). Accessed 16 August 2004.

Dobson, A. P., J. P. Rodriguez, W. M. Roberts, and D. S. Wilcove. 1997. "Geographic Distribution of Endangered Species in the United States." *Science* 275(5299): 550–553 (https://science.sciencemag.org/content/275/5299/550). Accessed 26 June 2019. https://doi.org/10.1126/science.275.5299.550

Dvorak, J. (ed.). 1993. *The PC Magazine. Computer Buyer's Guide*. Emeryville, CA: Ziff-Davis Press.

Endangered Wildlife Trust. 2010. *Annual Conservation Report of the Endangered Wildlife Trust* (https://www.ewt.org.za/conreport.html). Accessed 27 June 2017.

Environment Canada. *An Introduction to Climate Change – A Canadian Perspective* (http://www.msc-smc.ec.gc.ca/education/scienceofclimatechange/publications/reports_papers/index_e.html). Accessed 15 January 2007.

Environment Canada. *FAQ – Frequently Asked Questions CCD* (http://www.msc-smc.ec.gc.ca/education/scienceofclimatechange/publications/reports_papers/index_e.html). Accessed 15 January 2007.

Environment Canada. *Projections for Canada's Climate Future. CCD 00-01*. (http://www.msc-smc.ec.gc.ca/education/scienceofclimatechange/publications/reports_papers/index_e.html). Accessed 15 January 2007.

EUROPA. *A Quality Environment: How the EU is Contributing* (http://ec.europa.eu/publications/booklets/move/55/en.doc). Accessed 6 January 2007.

Field, S. *Introduction to STL, Standard Template Library* (http://www.linuxgazette.com/issue34/field.html). Accessed 16 August 2004.

Freedman, B., and D. E. McAllister. 2015. "Endangered Animals." *The Canadian Encyclopedia* (http://www.thecanadianencyclopedia.ca/en/article/endangered-animals/). Accessed 29 February 2016.

German Energy Agency (DENA). 2010. *Renewables - Made in Germany* (https://www.dena.de/fileadmin/dena/Dokumente/Pdf/9079_Flyer_Anbieterkatalog_renewables_Made_in_Germany_2016.pdf). Accessed 26 September 2019.

Gouvernement du Canada, Centre de recherches pour le développement international. "Case study: Brazil-Mercury contamination in the Amazon."

Gouvernement du Québec. 2012. "Transportation Electrification Strategy." *Priorité Emploi* (http://www.mce.gouv.qc.ca/publications/electrification-transports/strategie-electrification-en.pdf). Accessed 21 March 2014.

Hardesty, B. D., T. P. Good, and C. Wilcox. 2015. "Novel Methods, New results and Science-based Solutions to Tackle Marine Debris Impacts on Wildlife." *Ocean and Coastal management* 115 (http://www.sciencedirect.com/science/article/pii/S0964569115000897). Accessed 12 May 2017. https://doi.org/10.1016/j.ocecoaman.2015.04.004

Hruska, J. *Computer Virus Prevention: A Primer* (http://www.sophos.com/virusinfo/whitepapers/prevention.html). Accessed 13 October 2001.

Initiation au dépannage informatique (http://www.commentcamarche.net/repar/boot.php3). Accessed 24 May 2002.

Intergovernmental Panel on Climate Change (IPCC). 2006a. *An Assessment of Vulnerability*. (http://www.grida.no/climate/ipcc/spmpdf/region-e.pdf) Accessed 7 October 2006.

Intergovernmental Panel on Climate Change (IPCC). 2006b. *Climate Change 2001: The Scientific Basis (Summary for policymakers and technical summary)* (http://www.grida.no/climate/ipcc_tar/wg1/pdf/WG1_TAR-FRONT.PDF). Accessed 7 October 2006.

Intergovernmental Panel on Climate Change (IPCC). 2006c. *Climate Change and Biodiversity*. (http://www.ipcc.ch/pub/tpbiodiv.pdf). Accessed 7 October 2006.

Intergovernmental Panel on Climate Change (IPCC). 2006d. *Emissions Scenarios* (http://www.grida.no/climate/ipcc/spmpdf/sres-e.pdf). Accessed 7 October 2006.

Intergovernmental Panel on Climate Change (IPCC). 2006e. *Land Use, Land-Use Change, and Forestry* (http://www.grida.no/climate/ipcc/spmpdf/srl-e.pdf). Accessed 7 October 2006.

Intergovernmental Panel on Climate Change (IPCC). 2006f. *Safeguarding the Ozone Layer and the Global Climate System: Issues Related to Hydrofluorocarbons and Perfluorocarbons* (http://arch.rivm.nl/env/int/ipcc/pages_media/SROC-final/srocspmts_en.pdf). Accessed 7 October 2006.

An Internet Guide to Newcomers to the World Wide Web (http://www.about-the-web.com/). Accessed 14 August 2004.

Introduction to Computers (http://www.grassrootsdesign.com/intro/index.php). Accessed 4 August 2004.

Ivar do Sul, J. A., and Costa, M. F. 2014. "The Present and Future of Microplastic Pollution in the Marine Environment." *Environmental Pollution* 185: 352–364 (http://www.sciencedirect.com/science/article/pii/S0269749113005642). Accessed 18 May 2017. https://doi.org/10.1016/j.envpol.2013.10.036

Joshi, R. P., and A. P. Deshmukh. 2006. "Hybrid Electric Vehicles: The Next Generation Automobile Revolution." *Electric and Hybrid Vehicles ICEHV* (http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4156562). Accessed 2 April 2014.

Le Treut, J., G. Le Treut, and H. Le Treut. 2005. *Climate Change*. Paris: UNESCO Publishing.

Levesque, A., and J. A. Tremblay. 2008. *Rapport sur la situation de la chauve-souris rousse (Lasiurus borealis) au Québec* (ftp://ftp.mrn.gouv.qc.ca/Public/Bibliointer/Mono/2010/02/0998042.pdf). Accessed 19 November 2015.

Lindenmayer, D. B. et al. 2011. "How to Make a Common Species Rare: A Case against Conservation Complacency." *Conservation Biology* 144(5): 1663–1672. https://doi.org/10.1016/j.biocon.2011.02.022

Linux Directory. *A Beginner's Guide to Linux!* (http://www.linux-directory.com/newbie/). Accessed 4 August 2004.

LINUX Journal. Frequently Asked Questions (http://www.linuxjournal.com/modules.php?op=modload&name=FAQ&file=index&myfaq=yes&id_cat=2). Accessed 4 August 2004.

Meza-Parral, Y., and E. Pineda. 2015. "Amphibian Diversity and Threatened Species in a Severely Transformed Neotropical Region in Mexico." *PLoS One* 10(3). https://doi.org/10.1371/journal.pone.0121652

Mississippi Museum of Natural Science. 2014. *Endangered Species of Mississippi* (https://www.mdwfp.com/media/279708/endangered_species_of_mississippi.pdf). Accessed 10 February 2017.

Morrison, M. *Integrating Web Sites and Databases, Thomson Learning* (http://www.course.com/techtrends/morrison_102000.cfm). Accessed 13 October 2001.

Müller, S., and J. Buchdahl. 2000. *Encyclopedia of the Atmospheric Environment* (http://www.ace.mmu.ac.uk/eae/english.html). Accessed 07 January 2007.

National Ocean Service. 2007. "Welcome to Nonpoint Source Pollution." *NOS Education Discovery Kits* (http://oceanservice.noaa.gov/education/tutorial_pollution/lessons/pollution_tutorial.pdf). Accessed 10 May 2017.

New Mexico Department of Game and Fish. *Threatened and Endangered Species of New Mexico* (http://www.wildlife.state.nm.us/download/conservation/threatened-endangered-species/biennial-reviews/2014-Biennial-Review-Final.pdf). Accessed 14 February 2017.

Ontario Ministry of Energy and Infrastructure. 2010. Waterpower: An Overview (http://www.ontla.on.ca/library/repository/mon/24001/298188.pdf). Accessed 28 November 2019.

Ortiz, S. Jr. "Tips & Techniques For Using My Documents." *SmartComputing* (http://www.smartcomputing.com/editorial/article.asp?article=articles%2F2001%2Fs1211%2F06s11%2F06s11%2Easp&guid=wmhckule&searchtype=&WordList=). Accessed 1 November 2001.

PANACEA Environment English Monolingual Corpus (http://www.panacea-lr.eu/en/info-for-researchers/data-sets/monolingual-corpora/corpus-env-en/). Accessed 17 September 2015.

Paria, S. 2008. "Surfactant-enhanced Remediation of Organic Contaminated Soil and Water." *Advances in Colloid and Interface Science* 138(1): 24–58 (http://www.sciencedirect.com/science/article/pii/S0001868607001820). Accessed 15 May 2017. https://doi.org/10.1016/j.cis.2007.11.001

Parkanski, K. *How To Set Up Debian Linux Internet Servers* (http://www.aboutdebian.com/internet.htm). Accessed 16 August 2004.

Patz, J. T. *Canon Launches BJC Trio* (http://www.winmag.com). Accessed 14 October 1997.

Penn State College of Agricultural Science. 2014. "Pharmaceutical Disposal and Water Quality." (http://extension.psu.edu/natural-resources/water/drinking-water/water-testing/pollutants/pharmaceutical-disposal-and-water-quality). Accessed 15 May 2017.

Provost, G. "How Caching Works." *Howstuffworks* (http://www.howstuffworks.com/cache.htm). Accessed 5 January 2002.

Rehan, T. M. 1991. "Environmental Impact of Heavy Metal Pollution in Natural Aquatic Systems." PhD thesis presented at the University of West London (http://bura.brunel.ac.uk/bitstream/2438/5503/1/FulltextThesis.pdf). Accessed 12 May 2017.

Risley, D. *PC 101, PC Mechanic* (http://www.pcmech.com/show/internal/138/). Accessed 2 November 2001.

Rose, G. 2011. "E-bikes and Urban Transportation: Emerging Issues and Unresolved Questions." *Transportation* (http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=18&sid=3504d4b6-38c6-4f31-a52f-713a2712ba6f%40sessionmgr4001&hid=4209). Accessed 24 March 2014.

Savariar, V. 2014. "Environmental Health Monitoring: A Pragmatic Approach." *International Journal of Waste Resources* 4(4) (https://www.omicsonline.com/open-access/environmental-health-monitoring-a-pragmatic-approach-2252-5211-4-164.pdf). Accessed 12 May 2017.

Schwartz, P., and D. Randall. 2003. *Le scénario d'un brusque changement de climat et ses implications pour la sécurité nationale des Etats-Unis* (http://paxhumana.info/article.php3?id_article=427). Accessed 7 January 2007.

Shah, A., S. Shahzad, A. Munir, M. N. Nadagouda, G. S. Khan, D. F. Shams, D. D. Dionysiou, and U. A. Rana. 2016. "Micelles as Soils and Water Decontamination Agents." *Chemical Reviews* 116(10): 6042–6072 (http://pubs.acs.org/doi/full/10.1021/acs.chemrev.6b00132). Accessed 9 May 2017. https://doi.org/10.1021/acs.chemrev.6b00132

St-Pierre, A. 1988. *Le micro-ordinateur: introduction à MS/PC-DOS, LOTUS 1-2-3, dBASE III PLUS*. Boucherville: Éditions G. Vermette Inc.

Stohlgren, T. J., and K. Sunil. 2013. "Endangered Plants." *Encyclopedia of Biodiversity* 3: 205–215. https://doi.org/10.1016/B978-0-12-384719-5.00241-0

Swackhamer, D. L., H. W. Paerl, S. J. Eisenreich, J. Hurley, K. C. Hornbuckle, M. McLachlan, D. Mount, D. Muir, and D. Schindler. 2004. "Impacts of Atmospheric Pollutant on Aquatic Ecosystems" *Issues in Ecology* 12 (https://www.esa.org/esa/wp-content/uploads/2013/03/issue12.pdf). Accessed 12 May 2017.

Tate, E. D., M. O. Harpster, and P. J. Savagian. 2008. "The Electrification of the Automobile: From Conventional Hybrid, to Plug-in Hybrids, to Extended-Range Electric Vehicles." *Advanced Hybrid Vehicle Powertrain* (http://www.media.gm.com/content/dam/Media/microsites/product/volt/docs/paper.pdf). Accessed 4 April 2014.

Tsariounov, A. *Application note: Using LTT for System Analysis on Debian TE* (http://devresource.hp.com/drc/technical_papers/linuxltt.jsp [debian.txt] 5 195 mots). Accessed 4 August 2004.

Tyson, J. "How PCs Work." *Howstuffworks* (http://www.howstuffworks.com/pc.htm). Accessed 2 November 2001.

UICN. 2012. *Gorilles de Grauer et Chimpanzés de l'Est de la République Démocratique du Congo* (https://portals.iucn.org/library/efiles/documents/2012-080-Fr.pdf). Accessed 24 November 2015.

UNESCO. 2005. *The New Courrier*. Paris: United Nations Educational, Scientific and Cultural Organization (http://unesdoc.unesco.org/images/0014/001420/142021e.pdf#142040). Accessed 6 January 2007.

UNESCO. 2006. "A World of Science." *Natural Sciences Quarterly Newsletter* 4(1) (http://unesdoc.unesco.org/images/0014/001432/143225e.pdf). Accessed 7 October 2006.

U.S. Fish & Wildlife Service. 2008. *Endangered Species Bulletin* 33(3) (https://www.fws.gov/endangered/news/pdf/bulletin_fall2008.pdf). Accessed 15 June 2017.

Utah State University Extension Service Department of Fisheries and Wildlife. 1998. *Endangered and Threatened Animals of Utah* (http://www.berrymaninstitute.org/files/uploads/pdf/endgspec.pdf). Accessed 10 February 2017.

Whelton, A. J., I. McMillan, M. Connell, M. C. Kelley, J. P. Gill, K. D. White, and C. Novy. 2015. "Residential Tap Water Contamination Following the Freedom Industries Chemical Spill: Perceptions, Water Quality, and Health Impacts." *Environmental Science and Technology* 49: 813–823 (http://pubs.acs.org/doi/full/10.1021/es5040969). Accessed 9 May 2017. https://doi.org/10.1021/es5040969

"Why does a PC Crash?." *Henderson Computers* (http://www.hsncomp.co.nz/articles/ArticlesPCCrash.html). Accessed 21 December 2001.

Wildlife Services. 2001. "Effects of Wildlife Services on Predator Populations." *APHIS* (https://www.aphis.usda.gov/wildlife_damage/nwrc/downloads/effects_on_predators.pdf). Accessed 2 May 2017.

Wired Guide (http://www.wiredguide.com/wiredguide_training.htm). Accessed 5 August 2004.

Wołek, M., and O. Wyszomirski. 2013. "The Trolleybus as an Urban Means of Transport in the Light of the Trolley Project." *European Union* (http://www.zkmgdynia.pl/admin/__pliki__/TROLLEY_ebook_2013_.pdf). Accessed 14 April 2014.

World Alliance for the Decentralised Energy Association (WADE Thai). 2013. *Handbook on Smart/Intelligent Grid Systems Development and Deployment* (http://wadethai.org/docs/SMART_THAI_Handbook.pdf). Accessed 26 September 2019.

Yarsan, E., and M. Yipel. 2013. "The Important Terms of Marine Pollution 'Biomarkers and Biomonitoring, Bioaccumulation, Bioconcentration, Biomagnification'." *Journal of Molecular Biomarkers and Diagnosis* (https://www.omicsonline.org/the-important-terms-of-marine-pollution-biomarkers-and-biomonitoring-bioaccumulation-bioconcentration-biomagnification-2155-9929.S1-003.php?aid=11682). Accessed 12 May 2017. https://doi.org/10.4172/2155-9929.S1-003

Yukon Department of Environment (http://www.env.gov.yk.ca/). Accessed 27 June 2019.

Index

Lexical Semantics for Terminology: An introduction explores the interconnections between lexical semantics and terminology. More specifically, it shows how principles borrowed from lexico-semantic frameworks and methodologies derived from them can help understand terms and describe them in resources. It also explains how lexical analysis complements perspectives primarily focused on knowledge. Topics such as term identification, meaning, polysemy, relations between terms, and equivalence are discussed thoroughly and illustrated with examples taken from various fields of knowledge.

This book is an indispensable companion for those who are interested in words and work with specialized terms, e.g. terminologists, translators, lexicographers, corpus linguists. A background in terminology or lexical semantics is not required since all notions are defined and explained. This book complements other textbooks on terminology that do not focus on lexical semantics *per se*.

ISBN 978 90 272 0468 4
9 789027 204684

John Benjamins Publishing Company