languages

MDPI

Article

Phraseology and Culture in Terminological Knowledge Bases: The Case of Pollution and Environmental Law

Arianne Reimerink *, Pilar León-Araúz and Melania Cabezas-García

Department of Translation and Interpreting, University of Granada, 18071 Granada, Spain; pleon@ugr.es (P.L.-A.); melaniacabezas@ugr.es (M.C.-G.)
* Correspondence: arianne@ugr.es

Abstract: Despite its importance, environmental law has largely been ignored in environmental knowledge bases. This may be due to the fact that legal issues may not, strictly speaking, be considered scientific knowledge in environmental knowledge resources, which may in turn relate to the complexity of reflecting the cultural component (which includes different legal systems) in the description of terms and concepts. The terminological knowledge base EcoLexicon has recently begun to include information on environmental law. This paper takes the methodological perspective of frame-based terminology to analyze typical verb collocations in environmental law that will be added to the phraseology module of EcoLexicon. Corpus analysis was used to compare the behavior of verbs collocating with *pollution* in environmental science and environmental law. Verbs were classified based on lexical domains and semantic classes through definition factorization, as described in the Lexical Grammar Model. The differences were mostly based on the specificity of the arguments and the emphasis on the polluter in environmental law. This resulted in a proposal for the inclusion and configuration of environmental law phraseology in EcoLexicon, showing sociocultural differences across environmental subdomains.

Keywords: environmental law; terminological knowledge bases; phraseology; pollution

Citation: Reimerink, Arianne, Pilar León-Araúz, and Melania Cabezas-García. 2024. Phraseology and Culture in Terminological Knowledge Bases: The Case of Pollution and Environmental Law. *Languages* 9: 84. https://doi.org/10.3390/languages9030084

Academic Editor: Jeanine Treffers-Daller

Received: 1 December 2023
Revised: 8 February 2024
Accepted: 22 February 2024
Published: 29 February 2024

1. Introduction

Culture is generally regarded as the characteristics and knowledge of a particular group of people, encompassing religion, food, traditions, music, arts, and general language. As such, it permeates all aspects of life and even influences the way that we perceive the world (Unsworth et al. 2005). Not surprisingly, culture is also reflected in specialized language and terminology. Recently, the cultural facet of terminology or culture-bound terminology (Diki-Kidiri 2008) has been highlighted by Temmerman and van Campenhoudt (2014), Faber and Medina-Rull (2017), Diki-Kidiri (2022), Reimerink et al. (2023) and León-Araúz and Faber (Forthcoming). In fact, today, terms are acknowledged to possess an expressive power of their own insofar as they are often steeped in the culture and ideology of the text sender and even encode metaphors that have an impact on the understanding of a specialized domain (Faber 2022, p. 1). Since terms and their meanings are culturally motivated, the issue is how to represent this cultural dimension in terminological knowledge bases.

Recently, the process of converting EcoLexicon (ecolexicon.ugr.es) into an inclusive resource sensitive to cultural variation has driven the inclusion of new content and data categories. EcoLexicon is a multilingual and multimodal terminological knowledge base (TKB) (Faber et al. 2016) that represents the conceptual structure of the specialized domain of the environment in the form of a dynamic visual resource. It combines conceptual, linguistic, and graphical information to help translators, technical writers, and environmental experts acquire an in-depth understanding of specialized environmental concepts and help them write or translate specialized or semi-specialized texts. It is the practical application of

frame-based terminology (FBT) (Faber 2012, 2015, 2022), a cognitive approach to domain-specific language, which directly links specialized knowledge representation to cognitive linguistics and cognitive semantics. In FBT, knowledge acquisition begins at the term level, progresses to the phrase level, and finally results in the codification of an entire knowledge frame. The data are collected by means of corpus analysis.

To adapt EcoLexicon to cultural variation, a set of cultural profiles or frames must be specified that are linked to culture-dependent semantic categories, such as geographic landforms (e.g., *creek*), flora and fauna (e.g., *cookie-cutter shark*), meteorological phenomena (*local wind*), and even named entities (e.g., *Mesoamerican Reef System*). It also signifies adding a cultural component to all modules (definitions, conceptual networks, terms, phraseology, and multimodal resources). Culture in EcoLexicon is a broad notion that encompasses not only the inclusion of culture-specific concepts but also the different phraseological structures that arise from subtle changes in perspective (i.e., environmental subdomains) at the linguistic level.

Cultural variation is usually reflected in multidimensional concepts, whose relational behavior changes based on contextual parameters. Accordingly, cultural recontextualization depends on a set of cultural parameters, based on geographic location, historical time period, sociocultural usage, etc., which restrict the conceptual behavior to a certain cultural context. To reflect the sociocultural representation of environmental concepts, the information in EcoLexicon can be recontextualized according to environmental subdomain (e.g., geology, coastal engineering, hydrology, etc.). For example, the concept WATER has an active role in geology (it causes erosion, reshapes the terrestrial landscape, etc.), while in the water treatment domain, it is a patient that receives actions (purification, filtering, etc.) (León Araúz et al. 2013). An example of restrictions in conceptual networks for a concept that behaves differently according to its geographical location is WETLAND. In Figure 1, the network to the left shows the general network for WETLAND, whereas the network to the right is restricted to the Caribbean, with MARSH and SWAMP as prototypical wetlands for the area, and SEAGRASS BED, which is only there considered a wetland.

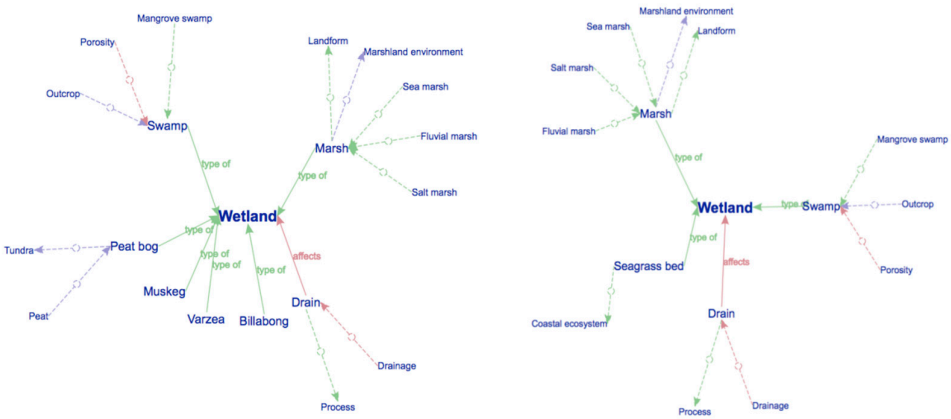

Figure 1. Non-restricted conceptual network (**left**) and network restricted for the Caribbean (**right**) for WETLAND.

Some subdomains, such as biodiversity, are more prone to cultural variation than others because flora and fauna are directly related to the geographical location they inhabit. However, there is one domain with a very special relationship to culture: environmental law. Environmental law is an important transversal domain that combines law with environmental science. It is impossible to understand the environment without an in-depth knowledge of how international, national, and regional governments and administrative bodies regulate it. The law is a profoundly human construct that is directly related to culture and, therefore, different in every culture. Studying the behavior of environmental concepts within this subdomain as compared to the environment as a whole promises to provide insight into the impact of culture on scientific knowledge. For this reason,

EcoLexicon has begun to include concepts and terms in different languages that pertain to environmental law (Faber and Reimerink 2019; Reimerink 2021).

In a previous study (Reimerink 2021), to expand and improve the information related to environmental law in EcoLexicon, comparative corpus analysis was used to identify missing concepts and explore how the multidimensional nature (León Araúz 2009) of environmental science might affect the behavior of other concepts in the subdomain of environmental law. The study focused on the POLLUTION frame, and the results showed that a new participant (i.e., the POLLUTER) had to be added when contextualized for the subdomain of environmental law. Whereas, in environmental science, the main focus is generally on the polluting substance, in environmental law, it is on the person/institution/industry responsible (see examples 1 and 2, emphasis by the authors). We also discovered that some facets of the concept POLLUTION (i.e., time and origin) are more prominent in this subdomain compared to the environmental domain as a whole (see examples 3 and 4).

1. The **pollutants** disperse in a downward direction, causing substantial air pollution at ground level but cannot escape upwards because of the inversion.
2. . . .the **polluter**- pays principle, the **person** responsible for the pollution cannot be identified or cannot be held liable under Community or national legislation. . .
3. Indeed, the phenomenon of **historical** pollution represents the result of the convergence and interaction of a number of different factors. . .
4. Historically the regulation of **vessel-source** pollution has engendered conflict between coastal States. . .

These results entailed changes in the conceptual networks and the definitions of EcoLexicon. Figure 2 shows the non-restricted conceptual network for POLLUTION without the generic–specific relations for more clarity.

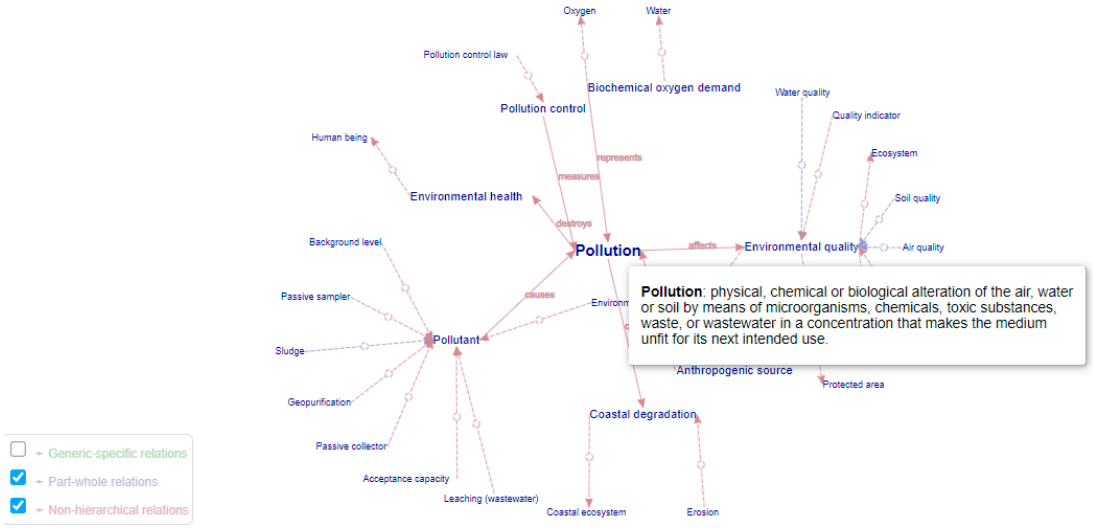

Figure 2. Conceptual network of POLLUTION without generic–specific relations and with a definition (Reimerink 2021).

Figure 3 shows the conceptual network for POLLUTION when applying contextual restrictions for the domain of environmental law. It includes the concept HISTORICAL POLLUTION, the additional participant POLLUTER, and the conceptual relations between the POLLUTER, the POLLUTANT, and POLLUTION.

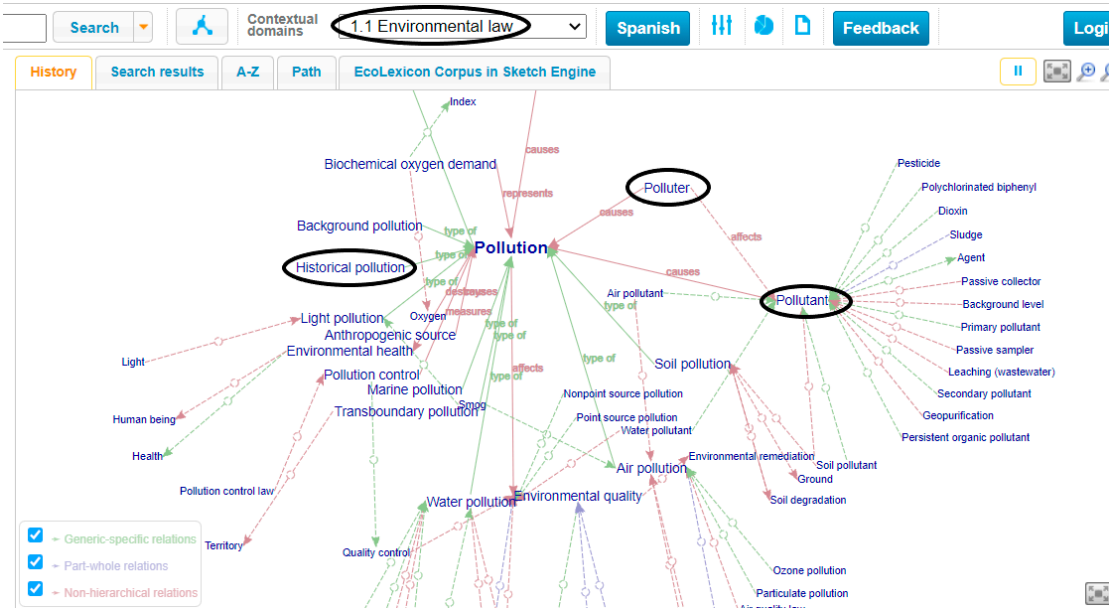

Figure 3. Conceptual network of POLLUTION recontextualized for the domain of environmental law (Reimerink 2021).

Although the final result in the conceptual network does not convey all the conceptual nuances, the relationship between POLLUTER, POLLUTANT, and POLLUTION is made explicit. The present case is a very good example of the need for multimodality in terminological knowledge bases. They must be enhanced with multimodal representations, namely visual and linguistic representations that converge to facilitate knowledge acquisition.

The results in Reimerink (2021) led to the revision of the definition of POLLUTION in EcoLexicon. A flexible definition was created to recontextualize it for environmental law. New facets included the facts that the polluter causes damage to the environment and that a polluter can be held responsible and sanctioned. The definitional template for POLLUTION (Table 1) now shows two agents. Agent1 is the polluter, who is ultimately responsible for the pollution. Agent2 is the pollutant, which is the direct cause of the pollution. The primary result (result1) is the direct consequence of pollution on the environment, whereas the secondary result (result2) is the fact that the polluter can be held responsible and sanctioned.

Table 1. Recontextualization of the definition of POLLUTION in environmental law.

POLLUTION (Environmental Law)	
Physical, chemical, or biological alteration of the air, water, or soil by means of microorganisms, chemicals, toxic substances, waste, or wastewater in a concentration that makes the medium unfit for its next intended use and that is caused by a person or company who can be held responsible under civil and/or criminal law	
type_of	PROCESS
has_agent1	PERSON/COMPANY
has_agent2	MICROORGANISM CHEMICAL TOXIC SUBSTANCE WASTE WASTEWATER
affects	AIR WATER SOIL
has_result1	UNFITNESS FOR INTENDED USE
has_result2	LIABILITY

Whereas the conceptual network provides graphical access to the POLLUTION frame and all the related concepts, including POLLUTER and POLLUTANT, the linguistic expression of the definition provides the means to convey the nuances of the relationship between the participants of the frame.

In the present study, we analyzed how the differences between environmental science and its subdomain environmental law, at the conceptual level, are conveyed at the linguistic level. End users of EcoLexicon, such as translators and technical writers, need to know how to express the differences at the conceptual level in their texts. This is usually reflected in phraseological combinations. However, even though the phraseology of specialized discourse is attracting increasing interest (Aguado de Cea 2007; Buendía-Castro 2013; Cabezas-García and Faber 2018), studies focusing on specialized phraseology are much less numerous than those addressing general language phraseology.

Our hypothesis is that the subdomain of environmental law uses different linguistic expressions to describe the POLLUTION frame than the global environmental science domain. The research questions we tried to answer are as follows: (a) how are the linguistic expressions related to the POLLUTION frame different when comparing environmental law and environmental science, and (b) how can we represent this knowledge in a TKB on the environment? The present study analyzed verb collocations in environmental law to add to the phraseology module of EcoLexicon, which is currently under construction. In this pilot study, we focus on phraseology in English. Future research will also address the topic in Spanish, one of the other major languages of EcoLexicon.

The rest of this paper is organized as follows: Section 2 explains the phraseology extraction method; Section 3 presents the results; Section 4 discusses the results and provides a proposal for their representation in the phraseology module; and Section 5 summarizes the conclusions that can be derived from this research.

2. Materials and Methods

In all cultures, legal language is a sublanguage with very specific syntactic, semantic, and pragmatic features (Tiersma 1999, pp. 15–133). The documents in the field often use grammatical structures that are rarely found elsewhere, such as redundancy, formulaic expressions, foreign words and Latinisms, syntactic discontinuity, impersonal and passive constructions, nominalization, and complex sentences (Hiltunen 2012; Williams 2004, pp. 112–15; Buendía-Castro and Faber 2015). Although, to a certain extent, the relation between content and form is present in other specialized texts as well, it is even more prevalent for texts in the legal domain since legal language is the result of a social contract and can be regarded as system-bound (Mattila 2006, p. 9).

Accordingly, an entry in a legal TKB can only be regarded as adequate if there is as complete a description as possible of the macro- and micro-context in which the term appears. If the resource is aimed at translators, for example, this description must provide information on how the term is used and the degree to which it can be regarded as equivalent to a given term within another legal system. Possible equivalent terms in other languages should also appear with as much contextual information as possible, which will facilitate mapping relations between the source and target language systems and cultures (Buendía-Castro and Faber 2015, p. 164). However, few specialized resources actually contain word combinations (L'Homme and Leroyer 2009, p. 260), and those that do include them are often not consistent in their treatment of phraseological units (Montero-Martínez and Buendía-Castro 2012).

Legal phraseology has attracted an increased interest in linguistics and translation studies. However, the same degree of interest has not been devoted to the issue of how phraseology can be managed and displayed in legal lexicographic and/or terminological resources (Peruzzo 2019, p. 149). In a questionnaire passed to final-year law students (Peruzzo 2019, p. 152), the students indicated that the enumeration of phraseological units in bi- or multilingual TKBs did not meet their needs because, firstly, these units were not accompanied by a definition and, secondly, in a bi- or multilingual terminological entry

containing a separate phraseology field for each term, establishing equivalence relations between phraseological units is not always a straightforward task.

The phraseology module of EcoLexicon is based on a wide interpretation of the concept of collocation, and at its core are verb collocations. An analysis of verb collocations in specialized discourse is especially relevant because they convey specialized knowledge and are essential to communicating fluently (Kübler and Pecman 2012; Orenha-Ottaiano et al. 2021; Buendía Castro 2021). In FBT, verb collocations are frequent combinations of two or more lexical units composed of a noun + verb, verb + noun, or noun + verb + noun, where the meaning of the verb is limited by the meaning of the noun. However, at the same time, the verb restricts the type of noun with which it can combine (Buendía-Castro 2013, p. 115). For example, in the collocation "the fire burns", the verb only allows for arguments that can be on fire, whereas the argument "fire" needs a verb that refers to the process of combustion (Montero Martínez and Buendía-Castro 2017).

In the phraseology module, verbs will be classified based on their meaning in combination with the terms with which they collocate. This is in line with previous work (Rosario et al. 2002; Maguire et al. 2010; Gagné and Spalding 2013; Cabezas García 2020), which analyzes the relevance of semantics in the recurrent patterns of combination that occur in phraseological units and the usefulness of these patterns in meaning access.

Therefore, verbs will not have their own entries in EcoLexicon but will be included as additional information in the term entries. The inclusion of a phraseme in EcoLexicon is essentially based on frequency of occurrence in the corpus. However, as will be shown, frequency changes when comparing different subdomains. Therefore, different phrasemes and examples will be shown, depending on the context the end user is focusing on in EcoLexicon.

To compare the collocational behavior of POLLUTION in environmental science and the subdomain of environmental law, Sketch Engine (https://www.sketchengine.eu/, Kilgarriff et al. 2014) was used. As a reference corpus, we used the EcoLexicon Environmental Corpus (EEC, 23 million words; León-Araúz et al. 2018) available in the Open Corpora section of Sketch Engine, and we compared it to a corpus specifically created for this purpose: the Environmental Law corpus (enLaw, 9.7 million words), composed of EEC texts, tagged with the domain of environmental law, as well as additional texts from the same domain harvested from the Internet. Some texts of the enLaw corpus are also included in the complete corpus on environmental science. Environmental law is part of the overall domain of environmental science; therefore, environmental law texts should also be included in the overall corpus. However, the differences between the overall domain as compared to the subdomain come to light when we compare the overall corpus with a corpus of texts that are specifically about environmental law. The EEC and enLaw corpora were both compiled in Sketch Engine with the Penn Treebank tagset and the EcoLexicon Semantic Sketch Grammar (ESSG; León-Araúz et al. 2016).

The ESSG is a Corpus Query Language (CQL)-based grammar (Jakubíček et al. 2013) as is the default grammar used for word sketches in Sketch Engine. Whereas Sketch Engine's default grammar provides grammatical relations, such as verb–object, modifiers, and prepositional phrases, the ESSG was developed for the extraction of semantic word sketches based on some of the most common semantic relations in terminology: generic–specific, part–whole, location, cause, and function. This was especially useful for the previous study (Reimerink 2021), where we focused on the conceptual differences between the global domain and the subdomain. However, to select representative examples for the phraseology module, the semantic word sketches provide easy access to sentences that convey conceptual knowledge (see Section 3, Figure 9). The Sketch Engine functions used to extract and compare the noun + verb collocations of *pollution*, as well as the related terms *pollute/polluter*, in both corpora were Word Sketch and Concordance.

After extraction, verbs were categorized according to the lexical domains in Faber and Mairal Usón (1999). The authors analyzed and categorized the semantic and syntactic structure of 12,000 general language English verbs through definition factorization, as described in the Lexical Grammar Model, and validated them via corpus analysis. This resulted in the following general lexical domains that can also be applied to verbs in specialized discourse: EXISTENCE (*be*, *happen*), CHANGE (*become*, *change*), POSSESSION (*have*), SPEECH (*say*, *talk*), EMOTION (*feel*), ACTION (*do*, *make*), MENTAL PERCEPTION (*know*, *think*), MOVEMENT (*move*, *go*, *come*), PHYSICAL PERCEPTION (*see*, *hear*, *taste*, *smell*, *touch*), MANIPULATION (*use*), CONTACT/IMPACT (*hit*, *break*), and POSITION (*put*, *be*). Other smaller classes included LIGHT, SOUND, BODY FUNCTIONS, WEATHER, etc.

3. Results

The results are presented according to the two functions of Sketch Engine used for corpus analysis: Word Sketch and Concordance.

3.1. Word Sketch

The information provided in Tables 2–6 is provided as Sketch Engine shows the data. The first column shows the collocate, the second column the absolute frequency, and the third the logDice score. The logDice score is used for determining how typical the collocation is. A high score means that the collocate is often found together with the node, and at the same time, there are not very many other nodes that the collocate combines with.[1]

Table 2. Word Sketch: first 25 verbs with *pollution* as object in enLaw and EEC.

enLaw			**EEC**		
Collocate	**Freq**	**Score**	**Collocate**	**Freq**	**Score**
2997	21.930		918	16.130	
control	299	10.940	air	27	9.670
cause	510	10.860	prevent	51	8.930
prevent	265	10.530	reduce	142	8.440
combat	138	10.300	control	44	8.380
reduce	243	10.120	cause	90	7.660
eliminate	95	9.740	minimize	13	7.580
air	41	8.760	combat	6	7.530
address	91	8.710	abate	5	7.450
regulate	58	8.530	eliminate	10	7.420
avoid	46	8.370	emit	12	7.300
minimise	30	8.170	avoid	10	6.960
abate	24	7.970	address	9	6.800
concern	60	7.880	regard	9	6.770
emit	25	7.860	create	18	6.650
limit	27	7.430	limit	11	6.410
regard	28	7.340	see	52	6.150
produce	26	7.310	increase	36	6.140
generate	20	7.250	indicate	11	5.960
tackle	15	7.200	generate	14	5.920
mitigate	14	7.040	monitor	5	5.920
minimize	14	7.020	associate	16	5.820
include	64	7.000	decrease	6	5.790
define	18	6.790	include	20	4.820
increase	22	6.720	produce	12	4.740
cover	15	6.520	consider	6	4.230

Table 3. Word Sketch: first 25 verbs with *pollution* as subject in enLaw and EEC.

enLaw			EEC		
Collocate	**Freq**	**Score**	**Collocate**	**Freq**	**Score**
1724	12.610		566	9.950	
cause	126	9.710	flush	18	9.830
affect	86	9.560	destroy	5	7.460
originate	34	9.150	affect	20	6.960
occur	46	8.680	reduce	8	6.660
arise	42	8.540	result	8	6.450
result	33	8.480	include	17	5.680
include	51	8.040	increase	8	5.670
pose	15	7.680	cause	13	5.650
damage	9	7.320	become	11	5.550
be	671	7.250	take	6	5.510
come	13	7.190	lead	5	5.390
emanate	8	7.150	occur	10	4.900
control	9	7.130	do	8	4.740
contribute	10	7.090	have	51	4.600
remain	14	7.060	be	226	4.030
derive	8	7.020			
permit	9	7.000			
impact	7	6.930			
continue	10	6.930			
threaten	7	6.830			
take	17	6.800			
follow	11	6.790			
harm	6	6.750			
have	142	6.650			
enter	7	6.570			

Table 4. Word Sketch: first 25 objects of *pollute* in enLaw and EEC.

enLaw			EEC		
Collocate	**Freq**	**Score**	**Collocate**	**Freq**	**Score**
1255	84.740		212	72.110	
activity	310	11.010	groundwater	14	9.560
industry	66	10.090	industry	8	8.640
substance	93	9.970	substance	8	7.720
matter	65	9.400	vehicle	3	7.320
facility	35	9.010	environment	11	7.310
discharge	27	8.930	river	4	7.210
firm	22	8.770	air	18	6.930
interference	19	8.690	stream	4	6.850
emission	46	8.670	gas	8	6.640
act	19	8.240	atmosphere	5	6.520
air	15	8.230	wastewater	2	6.420
behavior	14	8.140	supply	3	6.210
incident	12	8.050	emission	6	6.140
environment	35	8.000	km	2	5.910
factory	10	7.950	activity	4	5.820
company	15	7.890	water	19	5.780
behavior	12	7.880	fuel	2	5.750
effect	37	7.820	earth	3	5.550
water	24	7.760	behavior	2	5.550
event	10	7.740	product	3	5.510
conduct	10	7.680	step	2	5.280
technology	13	7.670	source	4	5.090
good	9	7.510	country	2	4.970
product	13	7.490	plant	3	4.930
process	20	7.390	beach	3	4.690

Table 5. Word Sketch: first 25 results for *polluter object_of* in enLaw and EEC.

enLaw			EEC		
Collocate	**Freq**	**Score**	**Collocate**	**Freq**	**Score**
346	21.750		13	13.270	
prosecute	20	10.270	divorce	1	11.090
sue	11	9.390	enshrine	1	10.410
deter	6	8.710	motivate	2	8.560
excuse	4	8.460	ascertain	1	8.540
oblige	10	8.430	hold	2	4.990
force	9	8.380	become	1	3.070
order	5	7.970	apply	1	2.550
compel	4	7.860	allow	1	2.440
let	3	7.760			
police	3	7.700			
pay	10	7.560			
allow	17	7.250			
get	3	7.040			
identify	12	7.030			
undermine	3	6.810			
locate	3	6.600			
find	6	6.350			
apply	6	6.230			
incorporate	3	6.220			
encourage	4	6.210			
bring	6	6.140			
require	21	6.030			
regulate	3	5.310			
implement	4	5.120			
see	3	4.710			

Table 6. Word Sketch: *polluter subject_of* in enLaw and EEC.

enLaw			EEC		
Collocate	**Freq**	**Score**	**Collocate**	**Freq**	**Score**
673	42.300		54	55.100	
pay	421	13.250	pay	47	12.540
bear	13	8.680	shape	1	6.760
violate	5	7.320	bear	1	6.490
cover	5	6.260	meet	1	5.420
contribute	3	6.190	provide	1	2.250
receive	3	6.040			
use	4	5.620			
cause	5	5.400			
take	3	4.650			
have	30	4.440			
be	93	4.410			
do	4	3.810			

Table 2 shows that the verbs that collocate with *pollution* as an object in both corpora mostly belong to the domain of CAUSATIVE EXISTENCE, more specifically to cause something to exist (*cause*), to cause something to cease to exist (*eliminate*), and to cause something to not happen (*prevent*, *avoid*). Other important lexical domains are CHANGE, more specifically, to cause something to change by decreasing it (*abate*, *reduce*, *minimize*, *mitigate*, *decrease*, *limit*) and MANIPULATION (*control*, *monitor*). Finally, the lexical domains of VISUAL PERCEPTION, COGNITION, and SPEECH are present with verbs such as *consider*, *define*, and *regard*.

In the word sketch of verbs with *pollution* as the subject (see Table 3), there are fewer results for the EEC because the numbers of collocations with *pollution* did not exceed the "auto" threshold, a default parameter in Sketch Engine based on corpus size.[2] This makes sense because the EEC is a corpus on the overall domain of environmental science; pollution is, thus, only one of the aspects to be considered. In contrast, in the enLaw corpus, POLLUTION is a central concept, and that is why collocations with *pollution* are statistically more relevant. The lexical domain of the verbs that predominate in both corpora is EXISTENCE: *originate*, *occur*, *arise*, *be*, *emanate*, *become*, and *include*. Another lexical domain present in both corpora is CHANGE (*reduce*, *increase*), to cause something to change by making it worse (*destroy*, *damage*, *harm*, *threaten*), and more general causative verbs such as *cause*, *affect*, *derive*, and *result*.

The verb *flush* in the EEC word sketch of *pollution* is the result of the term *pollution flushing*, which is a process through which pollution is removed from a water body through natural or artificial currents or tides. It can be classified as causing something to cease to exist (EXISTENCE) or as MOVEMENT (Faber and Reimerink 2019).

After analyzing *pollution*, we also analyzed the verb *pollute* and the noun *polluter* in Word Sketch. When we were looking at the results for the word sketch *object_of*, there were no obvious differences between the verb's behavior in enLaw and EEC, apart from the difference in the number of results (see Table 4). However, quite a few tagging mistakes were found, as some of the results are clearly objects (*air*, *environment*, *groundwater*, *river*, *beach*, *soil*, *surface*, *stream*, etc.), whereas others seemed to be clearly subjects of the verb (*industry*, *activity*, *discharge*, *emission*, *facility*, *behavior*, etc.). An example of the tagging mistakes is shown in the concordances for *polluting industries* in Figure 4, where *polluting* is obviously in an adjectival position. This shows that, although Word Sketch provides very valuable information in an easily accessible format, the processing of the corpora is not infallible, and therefore, manual analysis of concordances is necessary (see Section 3.2).

1	doc#0 ırks, especially since many are far from big cities and	**polluting**	**industries** . </s><s> However, air pollution carried f
2	doc#0 effect, which describes changes in the importance of	**polluting**	**industries** in a country once free trade is introduced
3	doc#0 nd low energy-efficiency in industry. </s><s> Heavily	**polluting**	energy **industry** . </s><s> Distorted pricing of goods
4	doc#0 nd low energy-efficiency in industry. </s><s> Heavily	**polluting**	energy **industry** . </s><s> Distorted pricing of goods
5	doc#0 son, 2000). </s><s> In the UK, increased taxation on	**polluting**	**industries** was a contentious issue for the Conserva
6	doc#0 urse, a difficulty with imposing new taxes or levies on	**polluting**	**industries** is that they will simply pass on such costs
7	doc#0 of administrative compensation paid for by levies on	**polluting**	**industry** seems the best fit with 'real' environmental
8	doc#0 onal standards for air quality, for example, will require	**polluting**	**industries** to invest a certain amount of money in pc
9	doc#0 om polluting activities (e.g. owners and employees of	**polluting**	**industries**) may experience a level of provision and

Figure 4. Concordance extract of *pollute* + *industry* available from word sketch *object_of pollute* in enLaw.

Table 5 shows which verbs collocate with *polluter* in the object slot. Once again, the enLaw corpus provides more results, some of which are directly related to the legal domain: *prosecute* and *sue*. This is why the concept of POLLUTER is only shown in relation to POLLUTION in the conceptual network restricted for environmental law. Another important lexical domain is MANIPULATION: *implement*, *regulate*, *oblige*, *force*, *compel*, *deter*, *require*, etc.

Finally, the word sketch *polluter subject_of* showed the verb *pay* as the very first result for both corpora. This is, of course, because one of the most important principles of environmental law is the polluter-pays principle (see Table 6).

3.2. Concordance

Apart from the fact that there were more results for *pollution* in enLaw, the lexical domains of the verbs collocating with *pollution* were very similar in both corpora. The differences pertained to some of the arguments of the verbs, which can be deduced from

the results of the Concordance function of Sketch Engine. To illustrate this, we analyzed the verbs (i) *abate* and (ii) *minimize*, both from the lexical domain CHANGE (to cause something to change by decreasing it), and (iii) *control* from the lexical domain MANIPULATION.

Figure 5 shows an extract of the concordances of the CQL *abate* + *pollution* in enLaw. The second argument that collocates with this combination is an institutional body (*state*, *UK*), a company (*industries*, *firms*), a measure (*measures*), or a cost (*expenditures*, *costs*).

1	doc#0	perate and reduces incentives for industries to	**abate**	**pollution** , for instance under the emissions trading scheme. <
2	doc#0	ed a common law suit brought by one state to	**abate**	**pollution** emanating from another state Tort law developed lar
3	doc#0	that those polluters that have the possibility to	**abate**	**pollution** above the regulatory standard do not, under a comm
4	doc#0	l control approach. </s><s> Some firms could	**abate**	**pollution** at relatively low costs or would innovate and invest ir
5	doc#0	take all appropriate measures to prevent and	**abate**	**pollution** caused by dumping from ships and aircraft (Article 5
6	doc#0	take all appropriate measures to prevent and	**abate**	**pollution** caused by dumping from ships and aircraft (Article 5
7	doc#0	al capital, and counts as GNP expenditures to	**abate**	**pollution** or to ameliorate environmental damage, giving a fals
8	doc#0	sense: each state should bear its own costs to	**abate**	**pollution** . -> The PPP in the wide sense: the polluting state sh
9	doc#0	sense: each state should bear its own costs to	**abate**	**pollution** . -- The PPP in the wide sense: the polluting state sh
10	doc#0	environment; the need to prevent, control, and	**abate**	**pollution** according to each state's capability; and particular sc
11	doc#0	lso the means by which the UK could not only	**abate**	**pollution** , but fulfil its obligations under EU law and other inte
12	doc#0	e 1992 Baltic Convention was to 'prevent and	**abate**	**pollution** and to protect and enhance the marine environment

Figure 5. Extract concordances *abate* + *pollution* in enLaw. Second arguments are highlighted with a rectangle.

Collocations of *abate* + *pollution* in the EEC corpus showed the same second arguments, which is not surprising, as all the occurrences were in texts tagged as pertaining to the environmental law domain or the water treatment domain.

The second argument for the CQL *minimize* + *pollution* (Figure 6) is mostly a measure (*requirements*, *directive*, *measures*) in enLaw.

1	doc#0	certification requirements for	**minimising** accidental	**pollution** ; and the carriage and discharge of oil-like substance
2	doc#0	irective 70/220/EEC was introduced to	**minimise** air	**pollution** from car exhaust fumes, and did this by prescribing l
3	doc#0	e offshore energy activities and attempt to	**minimise**	**pollution** from offshore installations. 9.1 INTRODUCTION: TH
4	doc#0	stal states to adopt measures designed to	**minimise**	**pollution** from installations in the seabed, and to protect and p
5	doc#0	es and to minimise waste; (c) prevent and	**minimise**	**pollution** of air, land and water in cost-effective ways; (d) incre
6	doc#0	er, the pilot or the salvor to prevent and/or	**minimise**	**pollution** .339 A s 137 direction may, for example, order the m
7	doc#0	waste within their area so as to prevent or	**minimise**	**pollution** of the environment or harm to human health. </s><s
8	doc#0	stem of IPC was established to prevent or	**minimise**	**pollution** of the environment due to the release of substances
9	doc#0	<s> In other words, the aim was clearly to	**minimise**	**pollution** by the application of BATNEEC, having regard to BP
10	doc#0	reasonably practicable steps to prevent or	**minimise**	**pollution** . </s><s> The onus is on the company or business s
11	doc#0	08/1/EC. </s><s> This directive aims at	**minimising**	**pollution** from various industrial sources throughout the EU. <

Figure 6. Extract concordances *minimise* + *pollution* in enLaw. Second arguments are highlighted with a rectangle.

However, the concordances for *minimize* + *pollution* in the EEC showed different second arguments (Figure 7). Infrastructural elements, such as water supply systems and wastewater treatment systems (concordance 2), locating wells in areas of deep groundwater and impermeable soils (4), bioethanol blending to petrol (6), the best available techniques not entailing excessive costs (9), and recycling techniques allied with good design practices (10) all refer to specific technical procedures developed by experts that have shown to be

the best options for minimizing pollution. Natural gas (7) and mangrove soils and plants (12), on the other hand, are natural entities that help minimize pollution.

1	<s>It also , possesses desirable features e. g. by **minimizing** the **pollution** near shore because it permits the flow exchange between the partially enclosed water body and the open sea.</s>
2	<s>Properly designed utilities and other infrastructural elements, such as water supply systems and wastewater treatment systems , **minimize pollution** and environmental damage.</s>
3	<s>The modern practice of burying electrical transmission lines in urban areas has **minimized** visual **pollution** from unsightly overhead power lines and made the power distribution system less vulnerable to interruptions from weather events.</s>
4	<s>The danger of **pollution** is **minimized** by locating wells in areas of deep ground water and impermeable soils</s>
5	<s>It also , possesses desirable features e. g. by **minimizing** the **pollution** near shore because it permits the flow exchange between the partially enclosed water body and the open sea.</s>
6	<s>If implemented bioethanol blending to petrol in Nepal could reduce Nepal 's fuel import which could save substantial money and could also **minimize** the environmental **pollution** .</s>
7	<s>The need for lowest cost and an ability to stockpile supplies favors coal , the need to **minimize** air **pollution** favors natural gas, and the need to be easily transported favors fuels derived from petroleum.</s>
8	<s>The NRSP defines " responsible environmental stewardship " as prudent use of all resources , **minimizing** unnecessary **pollution** , transforming waste into resources , and improving material conditions ; formulating practical environmental policies based on logic , scientific objectivity , and an understanding of risk ; individual rather than governmental action as the preferred means to achieve goals ; an understanding that private property encourages private responsibility ; a recognition that regulation of resources is best at the most local level possible ; and an understanding that more economic freedom allows more responsible individual action.</s>
9	<s>The Best Available Techniques Not Entailing Excessive Costs (BATNEEC) are required to **minimize pollution** of the environment as a whole.</s>
10	<s>The intelligent adoption of recycling techniques allied with good design practice which allows for materials to be reclaimed after their useful life is over , can do much to conserve raw materials and energy , **minimize pollution** and save money.</s>
11	<s>It provides useful information for devising strategies to **minimize pollution** and reduce the consumption of energy , water , and other resources. industrial ecosystem.</s>
12	<s>Although this is a new field of research there are strong and convergent indications that mangrove soils and plants **minimize pollution** by heavy metals in tropical coastlines.</s>
13	<s>It also **minimizes pollution** .</s>

Figure 7. Extract concordances *minimize* + *pollution* in EEC. Second arguments are highlighted with a rectangle.

The second argument for the CQL *control* + *pollution* (Figure 8) includes an institutional body (*state*, *administration*, *agencies*) and a measure (*strategies*, *measures*, *regulations*, *laws*) in enLaw. In the EEC, the second arguments fall in the same categories, again because the texts pertained to the environmental law, water treatment, and air quality management domains.

1	doc#0 lation is thus a system set up to prevent and **control**	**pollution**	. </s><s> A common distinction is that between tradi
2	doc#0 ıns (SIPs) that outline how each state will **control** air	**pollution**	under the Clean Air Act. </s><s> A SIP is a collectio
3	doc#0 </s><s> Article 17 Strategies to prevent and **control**	**pollution**	of groundwater 1. </s><s> The European Parliament
4	doc#0 ır In taking measures to prevent, reduce and **control**	**pollution**	of the marine environment, States shall act so as not
5	doc#0 laws and regulations to prevent, reduce and **control**	**pollution**	of the marine environment from activities in the Area
6	doc#0 uter may receive public subsidies for **controlling** the	**pollution**	. </s><s> In all these hypotheses, the affected comm
7	doc#0 ween the interests of the coastal states in **controlling**	**pollution**	in their waters, and the shipping interests of the flag s
8	doc#0 eas have usually entered into agreements to **control**	**pollution**	, but such regimes have been largely ineffectual give
9	doc#0 i> Separate environmental laws exist for **controlling**	**pollution**	in each medium. </s><s> In some cases, this compa
10	doc#0 to sue the administration for its failure to **control** the	**pollution**	, for example, Austria, Belgium, Germany, Spain, UK
11	doc#0 more by federal, state, and local agencies to **control**	**pollution**	, the use of environmental permit provisions by the g
12	doc#0 rules and standards to prevent, reduce and **control**	**pollution**	of the marine environment from vessels, and adopt r

Figure 8. Extract concordances *control* + *pollution* in enLaw. Second arguments are highlighted with a rectangle.

One of the participants that is specific to the pollution frame in environmental law is, evidently, the polluter. Figure 9 shows an extract of the concordances of the CQL *pollution caused_by* in enLaw. The cause is evidently the polluting industry (*ship*, *operational discharges*, *activities*) or the person or entity responsible (*polluters*, *manufacturers*, *persons*, *parties*, *corporation*). When choosing the examples for the phraseology module under the term *pollution*, the prominence of the polluter must be made explicit.

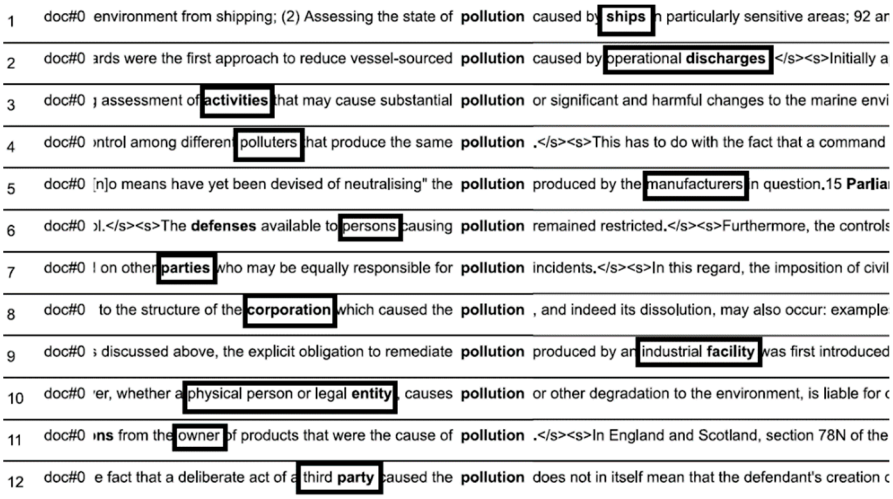

Figure 9. Extract concordances *pollution caused_by* in enLaw. Second arguments are highlighted with a rectangle.

4. Discussion

From the results shown in Section 3, certain conclusions can be drawn. First of all, POLLUTION is a much more central concept in the environmental law subdomain than in the general domain of environmental science. This can be deduced from the fact that, often, there were fewer results for the EEC than for enLaw, as the numbers of collocations with *pollution* in the EEC did not exceed the threshold, whereas in the enLaw corpus, the collocations with *pollution* were statistically more relevant.

Secondly, apart from the fact that there were more results for *pollution* in enLaw, the lexical domains of the verbs collocating with *pollution* were very similar in both corpora. The verbs that collocate with *pollution* as an object in both corpora mostly belonged to the domain of CAUSATIVE EXISTENCE, more specifically to cause something to exist, to cause something to cease to exist, and to cause something to not happen. Other important lexical domains were CHANGE, more specifically to cause something to change by decreasing it, and MANIPULATION. The word sketch of verbs with *pollution* as the subject showed that the lexical domain that predominates in both corpora is EXISTENCE. Another lexical domain present in both corpora is CHANGE.

When we were analyzing the verb *pollute* with the word sketch *object_of*, there were no obvious differences between the verb's behavior in enLaw and EEC. When we were studying the noun *polluter* as the object of verbs, verbs directly related to the legal domain such as *prosecute* and *sue* came up, and the most important lexical domain was MANIPULATION.

A few different second arguments arose when we were analyzing the concordances of the verbs *abate*, *minimize*, and *control*. Especially the categories for the second argument of *minimize* were very different in enLaw (measure) as compared to the EEC (technical procedures and natural entities).

Regarding the phraseology module in EcoLexicon, the verbs *abate*, *minimize*, and *control* will be included under the term *pollution* in the following phrasemes for the environmental law subdomain:

- INSTITUTIONAL BODY/COMPANY/MEASURE/COST + CHANGE [decrease: *abate*, *minimize*] + POLLUTION

- INSTITUTIONAL BODY/MEASURE + MANIPULATION [*control*] + POLLUTION

The first phraseme for the environmental science domain as a whole will be different:

- INSTITUTIONAL BODY/COMPANY/MEASURE/COST/TECHNICAL PROCEDURE/NATURAL ENTITY + CHANGE [decrease] + POLLUTION

The examples of collocations for the phraseology module will be chosen so as to highlight the differences between the arguments in environmental law and environmental science, showing different examples, depending on the contextualization of the POLLUTION frame.

Table 7 shows the information that will be included in EcoLexicon's phraseology module. Under the term *pollution* within the subdomain of environmental law, the different lexical domains will be presented with the verbs identified by corpus analysis. When clicking on each verb, the second argument categories will be shown (e.g., INDUSTRY, INSTITUTIONAL BODY, and PERSON/COMPANY in the first row [EXISTENCE, cause to exist]). When clicked on, example sentences that illustrate these verbs and arguments will also appear. In the table, the example sentences are shown for INSTITUTIONAL BODY *causes* POLLUTION, PERSON/COMPANY *causes* POLLUTION, INSTITUTIONAL BODY *abates* POLLUTION, COMPANY *abates* POLLUTION, and MEASURE *abates* POLLUTION.

Table 7. Proposal for phraseology module related to the term *pollution* in EcoLexicon.

Pollution [Environmental Law]			
EXISTENCE [cause to exist]	cause	INDUSTRY	
		INSTITUTIONAL BODY	On the other hand, if state B causes pollution in state A, state A is entitled to invoke its territorial sovereignty
		PERSON/COMPANY	The cost is borne by the company who causes the pollution or transferred to consumers driving demand for the relevant product
EXISTENCE [cause to not exist]	eliminate tackle		
CHANGE [decrease]	abate decrease limit minimize mitigate reduce	INSTITUTIONAL BODY	Courts have allowed a common lawsuit brought by one state to abate pollution emanating from another state
		COMPANY	Some firms could abate pollution at relatively low costs
		MEASURE	The 1976 BARCON requires parties to take all appropriate measures to prevent and abate pollution caused by dumping from ships and aircraft
		COST	
MANIPULATION	combat		
	control monitor		

As the examples show, emphasis in the environmental law domain is on the polluter (e.g., "state B causes pollution", "The cost is borne by the company who causes pollution"), the liability of the polluter before the courts ("Courts have allowed a common law suit. . ."), and the facets of the POLLUTION frame that stand out in environmental law: time and

origin ("...dumping from ships and aircraft"). However, if the phraseology for POLLUTION is contextualized for environmental science, the phraseme for CHANGE [decrease] will change: the categories TECHNICAL PROCEDURE and NATURAL ENTITY will be shown, and the example sentences will change their focus to the polluting substance.

5. Conclusions

The results described in this paper show that frame-based terminology provides the methodological underpinnings to extract the subtle differences between environmental science and its subdomains at the linguistic level. Specifically, verbal collocations in the environmental law domain differ from those in the environmental science domain in regard to the specificity of the arguments or even the activation of certain verbs. These differences must be included in terminological knowledge bases in order to provide an accurate representation of environmental knowledge, as they reveal the nuanced ways in which language is used across different contexts to discuss similar issues. For example, in the broader environmental domain, verbs associated with POLLUTION might include general actions like *reduce*, *prevent*, and *control*, reflecting a wide range of activities impacting the environment. Conversely, within the subdomain of environmental law, the phraseology becomes more precise, incorporating legal-specific verbs such as *regulate* and *sue*. This shift in terminology not only underscores the importance of context-specific language for clarity and precision in discourse but also highlights a conceptual change of perspective. Differences at the conceptual level pervade the linguistic level because of the choice of verbs and their arguments. In the same way, the differences observed at the linguistic level can contextualize the conceptual representation of specialized concepts in the conceptual networks.

The present study adds to the still scarce research in specialized phraseology, as well as studies in legal phraseology, which, to our knowledge, have not touched upon legal phraseology in scientific domains. Furthermore, it provides a proposal as to how to represent this phraseology in a terminological resource. The representation proposal, where the verbs of the phraseme are classified according to the lexical domain and the arguments are classified in broader semantic categories, provides a direct link between the phraseme and its underlying semantics. It, therefore, provides the necessary knowledge for end users when they need to choose between different phraseological options.

Representing this phraseological knowledge for all the terms in EcoLexicon in English and in Spanish will be one of the challenges for the future development of EcoLexicon.

Author Contributions: Conceptualization, A.R. and P.L.-A.; methodology, A.R. and M.C.-G.; formal analysis, A.R.; investigation, A.R., M.C.-G. and P.L.-A.; writing—original draft preparation, A.R.; writing—review and editing, M.C.-G. and P.L.-A.; funding acquisition, P.L.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was carried out under the project TRANSCULTURE, reference number PID2020-118369GBI00, funded by the Spanish Ministry of Science and Innovation, MCIN/AEI/10.13039/501100011033.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The EEC corpus analyzed in this study is publicly available at the Open Corpora section of Sketch Engine: https://app.sketchengine.eu. The enLaw corpus is available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Notes

[1] For more information on statistics in Sketch Engine: https://www.sketchengine.eu/wp-content/uploads/ske-statistics.pdf.

[2] For more information: https://www.sketchengine.eu/documentation/methods-documentation/ (accessed on 21 February 2024).

References

Aguado de Cea, Guadalupe. 2007. La fraseología en las lenguas especializadas. In *Las Lenguas Profesionales y Académicas*. Edited by Enrique Alcaraz Varó, José Mateo Martínez and Francisco Yus Ramos. Madrid: Ariel, pp. 53–65.

Buendía-Castro, Míriam. 2013. Phraseology in Specialized Language and Its Representation in Environmental Knowledge Resources. Ph.D. thesis, Universidad de Granada, Granada, Spain.

Buendía Castro, Míriam. 2021. *Verb Collocations in Dictionaries and Corpus: An Integrated Approach for Translation Purposes*. Berlin: Peter Lang.

Buendía-Castro, Míriam, and Pamela Faber. 2015. Phraseological units in English-Spanish legal dictionaries: A comparative study. *Fachsprache: International Journal of Specialized Communication* XXXVII: 161–75.

Cabezas García, Melania. 2020. *Los términos compuestos desde la Terminología y la Traducción*. Berlin: Peter Lang. [CrossRef]

Cabezas-García, Melania, and Pamela Faber. 2018. Phraseology in specialized resources: An approach to complex nominals. *Lexicography* 5: 55–83. [CrossRef]

Diki-Kidiri, Marcel. 2008. *Le Vocabulaire Scientifique dans les Langues Africaines. Pour Une Approche Culturelle de la Terminologie*. Paris: Karthala.

Diki-Kidiri, Marcel. 2022. Cultural Terminology. An introduction to theory and method. In *Theoretical Perspectives on Terminology: Explaining Terms, Concepts and Specialized Knowledge*. Edited by Pamela Faber and Marie-Claude L'Homme. Amsterdam: John Benjamins, pp. 197–216. [CrossRef]

Faber, Pamela. 2012. *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin and Boston: De Gruyter Mouton.

Faber, Pamela. 2015. Frames as a Framework for Terminology. In *Handbook of Terminology*. Edited by Hendrik J. Kockaert and Frieda Steurs. Amsterdam: John Benjamins, vol. 1, pp. 14–33.

Faber, Pamela. 2022. Frame-based Terminology. In *Theoretical Perspectives on Terminology: Explaining Terms, Concepts and Specialized Knowledge*. Edited by Pamela Faber and Marie-Claude L'Homme. Amsterdam: John Benjamins, pp. 353–76. [CrossRef]

Faber, Pamela, and Ricardo Mairal Usón. 1999. *Constructing a Lexicon of English Verbs*. Berlin: Mouton de Gruyter.

Faber, Pamela, and Laura Medina-Rull. 2017. Written in the Wind: Cultural Variation in Terminology. In *Cognitive Approaches to Specialist Languages*. Edited by Marcin Gryviel. Newcastle-upon-Tyne: Cambridge Scholars, pp. 419–42.

Faber, Pamela, and Arianne Reimerink. 2019. Framing Terminology in Legal Translation. *International Journal of Legal Discourse* 4: 15–46. [CrossRef]

Faber, Pamela, Pilar León-Araúz, and Arianne 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, Portoroz, Slovenia*. Edited by Ilan Kernerman, Iztok Kosem Trojina, Simon Krek and Lars Trap-Jensen. pp. 73–80. Available online: http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-GLOBALEX_Proceedings-v2.pdf (accessed on 21 February 2024).

Gagné, Christina L., and Thomas L. Spalding. 2013. Conceptual composition: The role of relational competition in the comprehension of modifier-noun phrases and noun-noun compounds. *Psychology of Learning and Motivation* 59: 97–130.

Hiltunen, Risto. 2012. The Grammar and Structure of Legal Texts. In *The Oxford Handbook of Language and Law*. Edited by Lawrence M. Solan and Peter M. Tiersma. Oxford: Oxford University Press, pp. 39–51.

Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý, and Vít Suchomel. 2013. The TenTen Corpus Family. Paper presented at the 7th International Corpus Linguistics Conference CL 2013, Lancaster, UK, July 22–26; pp. 125–127.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. The Sketch Engine: Ten Years on. *Lexicography* 1: 7–36. [CrossRef]

Kübler, Natalie, and Mojca Pecman. 2012. The ARTES bilingual LSP dictionary: From collocation to higher order phraseology. In *Electronic Lexicography*. Edited by Sylviane Granger and Magali Paquot. Oxford: Oxford University Press, pp. 187–209.

León Araúz, Pilar. 2009. Representación Multidimensional del Conocimiento Especializado: El Uso de Marcos desde la Macroestructura hasta la Microestructura. Ph.D. thesis, Universidad de Granada, Granada, Spain.

León Araúz, Pilar, Arianne Reimerink, and Alejandro García Aragón. 2013. Dynamism and context in specialized knowledge. *Terminology* 19: 31–61. [CrossRef]

León-Araúz, Pilar, Antonio San Martín, and Pamela Faber. 2016. Pattern-based Word Sketches for the Extraction of Semantic Relations. Paper presented at the 5th International Workshop on Computational Terminology (Computerm2016),COLING 2016, Osaka, Japan, December 12; pp. 73–82.

León-Araúz, Pilar, Antonio San Martín, and Arianne Reimerink. 2018. The EcoLexicon English Corpus as an open corpus in Sketch Engine. Paper presented at the 18th EURALEX International Congress, Ljubljana, Slovenia, July 17–21; Edited by Jaka Čibej, Vojko Gorjanc, Iztok Kosem and Simon Krek. Ljubljana: Ljubljana University Press, pp. 893–902.

León-Araúz, Pilar, and Pamela Faber. Forthcoming. Including the cultural dimension of terminology in a fram-based resource. In *Terminology and Cognition*. Berlin: Mouton de Gruyter.

L'Homme, Marie-Claude, and Patrick Leroyer. 2009. Combining the semantics of collocations with situation driven search paths in specialized dictionaries. *Terminology* 15: 258–83. [CrossRef]

Maguire, Phil, Edward J. Wisniewski, and Gert Storms. 2010. A corpus study of semantic patterns in compounding. *Corpus Linguistics and Linguistic Theory* 6: 49–73. [CrossRef]

Mattila, Heikki E. S. 2006. *Comparative Legal Linguistics*. Aldershot: Ashgate.

Montero Martínez, Silvia, and Míriam Buendía-Castro. 2017. Clasificación semántica de colocaciones verbales para la adquisición y codificación de conocimiento experto: El caso de los riesgos naturales. *Revista Española de Lingüística Aplicada* 30: 240–72. [CrossRef]

Montero-Martínez, Silvia, and Míriam Buendía-Castro. 2012. La sistematización en el tratamiento de las construcciones fraseológicas: El caso del medio ambiente. In *Empiricism and Analytical Tools for 21st Century Applied Linguistics*. Edited by Izaskun Elorza, Ovidi Carbonell i Cortés, Reyes Albarrán, Blanca García Riaza and Miriam Pérez-Veneros. Salamanca: Ediciones de la Universidad de Salamanca, pp. 711–24.

Orenha-Ottaiano, Adriane, Marcos García, Maria Eugênia Olímpio de Oliveira Silva, Marie-Claude L'Homme, Margarita Alonso Ramos, Carlos Roberto Valêncio, and William Tenório. 2021. Corpus-based Methodology for an Online Multilingual Collocations Dictionary: First Steps. In *Electronic Lexicography in the 21st Century, Proceedings of the eLex 2021 Conference, Virtual, 5–7 July 2021*. Brno: Lexical Computing CZ, s.r.o., pp. 1–28.

Peruzzo, Katia. 2019. Developing targeted legal terminology resources: Learning from future lawyers. In *New Challenges for Research on Language for Special Purposes*. Edited by Ingrid Simonnæs, Øivin Andersen and Klaus Schubert. Berlin: Frank & Timme, pp. 141–58.

Reimerink, Arianne. 2021. Pollution in Environmental Law: Comparative Corpus Analysis. *International Journal of Lexicography* 35: 204–233. [CrossRef]

Reimerink, Arianne, Pamela Faber, Melania Cabezas-García, and Pilar León-Araúz. 2023. Legal Jargon in an Environmental TKB: Pollution Phraseology. Paper presented at the 2nd International Conference on "Multilingual Digital Terminology Today. Design, Representation Formats and Management Systems" (MDTT) 2023, Lisbon, Portugal, June 29–30; pp. 29–30.

Rosario, Barbara, Marti Hearst, and Charles Fillmore. 2002. The descent of hierarchy, and selection in Relational Semantics. Paper presented at the 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, PA, USA, July 7–12; Philadelphia: ACL, pp. 247–54.

Temmerman, Rita, and Marc van Campenhoudt. 2014. *Dynamics and Terminology*. Amsterdam: John Benjamins.

Tiersma, Peter. 1999. *Legal Language*. Chicago: University of Chicago Press.

Unsworth, Sara J., Christopher R. Sears, and Penny M. Pexman. 2005. Cultural Influences on Categorization Processes. *Journal of Cross-Cultural Psychology* 36: 662–88. [CrossRef]

Williams, Christopher. 2004. Legal English and Plain Language: An introduction. *ESP Across Cultures* 1: 111–24.