Terminological Harmonization of Environmental Law Principles

Arianne Reimerink[1, †,*] and Pilar León-Araúz[1, †]

[1] University of Granada, Departamento de Traducción e Interpretación, Buensuceso 11, 18071 Granada, Spain

Abstract

The EU plays an important role in legal harmonization. In this study we analyze if legal harmonization also entails terminological harmonization, as terminological consistency is essential in international organizations. A diachronic analysis was carried out of environmental law principles and their term variants in English and Spanish in the EUR-Lex2/2016 corpus. The results show a high degree of term variation in both languages and only a slight tendency towards harmonization where the use of certain variants increases over time and the number of different variants for the same principle decreases. The complexity of the multiword terms studied is probably one of the causes for the astonishing number of variants found, but we believe that more effort and resources should be allocated to term management.

Keywords

Terminological harmonization, environmental law principles, translation, diachronic corpus analysis

1. Introduction

Over the years, Environmental Law has developed thanks to initiatives at the national, regional and international levels. Principles, such as the precautionary principle, prevention principle, and polluter pays principle, are now fully integrated into environmental law at an international, regional and, often, national level. Some of these principles originated at the national level, and have permeated into regional and international law. In contrast, others, which were initially implemented at a regional or international level, have later on been adopted by national lawmakers [1].

From the 1970s onwards, the EU has played an important role in the harmonization of national environmental laws. The harmonization of law can be defined as ensuring concordance of various legal orders through the elimination of divergent elements and overcoming differences so that these orders function in accordance with the aims and interests of the engaged entities [2].

In the present study, we assume that legal harmonization will (and should) in time also lead to terminological harmonization. As all EU regulations and directives are made available in all the official languages of the EU, terminological harmonization should occur in the translation process as well. Legal terms, the main "prompts and points of access to knowledge structures of the domain" [3], constitute a central feature of legal translation [4]. Furthermore, ensuring terminological consistency is essential in international organizations for the sake of legal univocity and certainty, and thus for translation quality assurance [5].

To study terminological harmonization in environmental law, we carried out a diachronic analysis of the terms used for the principles of environmental law in the EUR-Lex2/2016 corpus (EUR-Lex) [6] available in Sketch Engine [7]. As will be seen, the type of variation studied here was denominative variation, where different designations are used to name the same concept [8], as opposed to conceptual variation, which occurs when concepts can be organized according to

* Corresponding author.
† These authors contributed equally.
✉ arianne@ugr.es (A. Reimerink); pleon@ugr.es (P. León-Araúz)
ⓘD 0000-0002-7264-4580 (A. Reimerink); 0000-0002-8520-2749 (P. León-Araúz);

different facets or dimensions [9]. The rest of the paper is organized as follows. Section 2 explains the method used for analysis. In section 3, the results are shown and discussed. Finally, in section 4 some conclusions are drawn and ideas for future research are presented.

2. Method

The English-Spanish parallel subcorpora of the EUR-Lex2/2016 corpus were used to study terminological harmonization of environmental law principles. The corpus is subdivided diachronically into decades, thus providing the means to analyze changes over time. We analyzed how the most salient environmental law principles are conveyed in English (50s, 70s, 80s, 90s, 2000s, and 2010s) and Spanish (80s, 90s, 2000s, 2010s; data for Spanish are available from 1986 onwards).

Firstly, to make sure that the corpus we analyzed was related to environmental law (the topic of this study), we used the option EUROVOC classification in Text types of the function Concordance. EUROVOC is the EU's multilingual and multidisciplinary thesaurus, which contains keywords organized in 21 domains and 127 sub-domains. In this case, we restricted the texts in the corpus to "all values containing environment". In a first run, we analyzed the complete English subcorpus to select the environmental law principles, whose diachronical evolution we would study afterwards, with the following CQL:

CQL1: [tag="N.*|JJ.*|RB.*|VVN.*|VVG.*"]{1,}[lemma="principle"][tag!="N.*|JJ.*"] within <s/>

It searches, within a sentence (within <s/>), for the lemma *principle* ([lemma="principle"]) preceded by nouns, adjectives, adverbs, past participles, or present participles ([tag="N.*|JJ.*|RB.*|VVN.*|VVG.*"]) appearing one or more times ({1,}) and followed by anything except for a noun or adjective. This prevents from extracting sentences where principle is not the head of the term but a modifier. The search provided 4722 results, which were manually analyzed to extract MWTs such as *precautionary principle*, *prevention principle*, and *polluter pays principle*, as well as several of their variants.

After selecting the environmental principles and variants, these were searched in the subcorpora divided by decade and with the same EUROVOC classification restriction. At this point, two additional CQL searches where carried out to find further variants:

CQL2:
[lemma="principle"][]{0,5}[lemma="polluter|pay|prevention|preventive|sustain|precaution|precautionary|rectification|rectify|remedy|source"] within <s/>

CQL3:
[lemma="polluter|pay|prevention|preventive|sustain|precaution|precautionary|rectification|rectify|remedy|source"][]{0,5}[lemma="principle"] within <s/>

CQL2 and CQL3 search for possible variants by using the MWT head "principle" and the modifiers found in the previous search (for example *polluter*, *prevention*, *source*, etc.) with a 5-element span in between. This way, variants such as "principle that the polluter should pay" could be extracted.

Finally, the Parallel Concordance function of Sketch Engine was used to analyze term variation in the Spanish version of the English concordances. This provided information on whether certain English variants show a preferred Spanish variant, and if those preferences also change over time and become more harmonized.

3. Results and discussion

3.1. Preliminary corpus analysis

In the first run, where we analyzed the complete corpus, CQL1 retrieved 4722 concordances. As the CQL sequence may cause several hits for each concordance[2], after hiding sub-hits in the advanced Filter function, 3950 unique results were retrieved. These were analyzed manually to get a feel for the terms and variants related to the environmental law principles and the language surrounding them in the corpus. Concordances such as the following were useful for our analysis:

1. The 7th EAP should therefore be based on the fundamental principles of environmental law – the polluter pays principle, the precautionary principle, the preventive principle and the rectification at source principle – …
2. … improved implementation of key environmental principles (polluter pays, prevention and precaution) and of existing EC environmental laws …
3. The Contracting Parties undertake to develop the transport sector while observing the precautionary principle, the preventive principle and the polluter-pays principle.
4. In addition, this regulation is consistent with the Union's environmental legislation and policy and its main tenets such as pollution prevention, control and the polluter pays and precautionary principles.

In 1 and 2 the keywords in context (KWIC) were "fundamental principles" and "environmental principles" respectively, whereas in 3 it was "precautionary principle" and in 4 "precautionary principles".

Other results were not taken into account as they were not related to environmental principles, as shown in the following examples:

5. … the basic principles for integrated protection and use of groundwaters and surface waters within such a river basin management approach …
6. … budgetary principles of the Member States, such as unity and universality, should be used to reduce greenhouse gas emissions …
7. … should be guided by these climate policy principles and by the principles of sustainability, social responsibility and equity between the generations and people …

Although example 5 refers to basic principles related to protection, the context clarifies that these principles are more related to the scientific and technical principles that can be applied and not the legal principles. Example 6 refers to another subfield (budget) as well and 7 distinguishes between climate policy principles and other, thus not related, principles.

This first run also gave a first insight into the high level of term variation of some of the principles involved. For example, for the polluter pays principle the following were found: *polluter pays principle, polluter-pays principle, "polluter pays" principle, 'Polluter Pays' principle, "polluter pays principle", principle that the polluter should pay*, etc.

After this first run, the following principles stood out for further analysis: the polluter pays principle, the precautionary principle, the prevention principle, the subsidiarity principle, the sustainability principle, and the rectification at source principle (not the preferred variant as we will see later on).

3.2. Diachronic variant analysis in English

For the second run, we analyzed the selected principles in the subcorpora restricted according to decade with the additional restriction of "all values containing environment" in the EUROVOC

[2] Queries in Sketch Engine are both lazy and greedy.

option of Text Types. Again sub-hits were omitted. No results were obtained for the 1950s and 1970s, which were thus excluded from further analysis, and only very few for the 1980s. The subsidiarity principle and sustainability principle were also excluded for lack of results, 10 and 8 hits respectively over all decades in total. Variants for each principle were extracted and counted manually.

Two limitations must be taken into account: 1) the number of tokens in each subcorpus is different; and 2) the 2010s only include texts up to 2016. Therefore, the results and our conclusions must be taken with caution and will mostly refer to tendencies that stand out according to the data we have.

For the polluter pays principle (Figure 1), the variants *'polluter pays' principle* and *polluter pays principle* stand out in the 1990s, 2000s, and 2010s. The variant *principle that the polluter should pay* is quite common and stays stable over time. However, the *"polluter pays" principle* is the second most preferred variant in the 2000s (122 hits) but appears much less often in the 2010s (9). There seems to be a tendency of terminological harmonization from the 2000s to the 2010s as quite a few variants disappear (20), although a few new ones come up (4). Taking into account that the 2010s has fewer tokens and only includes texts up to 2016, no definitive conclusions can be drawn.
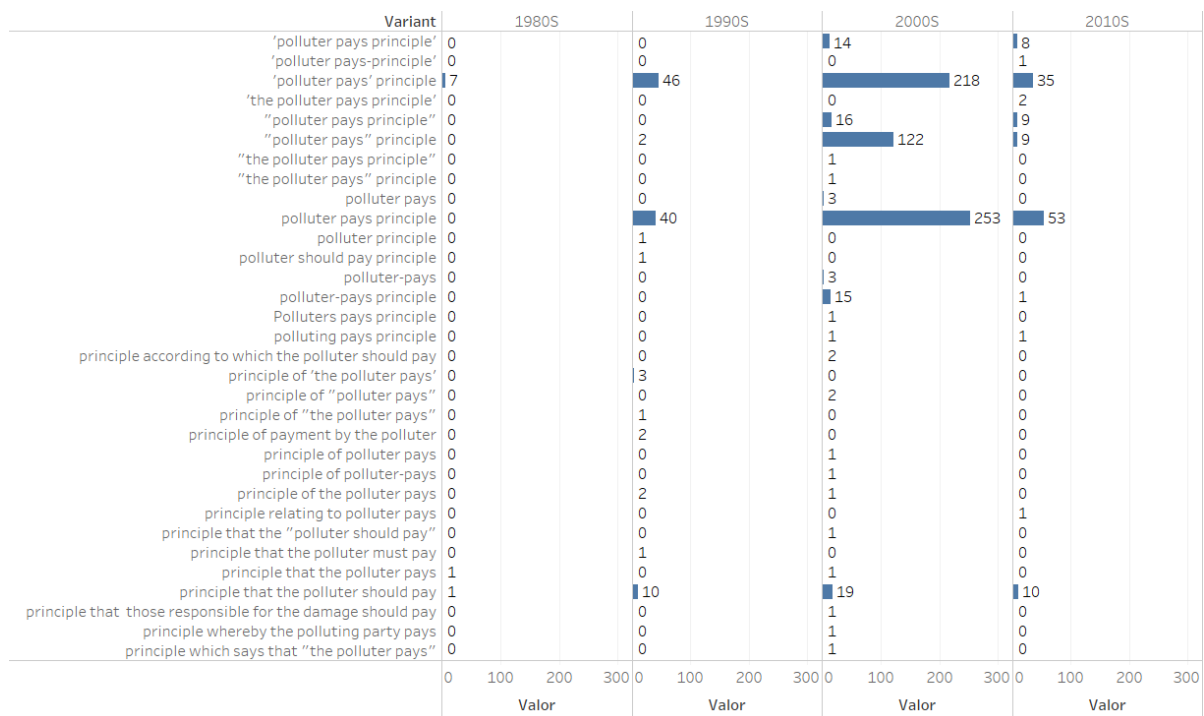
Variant	1980S	1990S	2000S	2010S
'polluter pays principle'	0	0	14	8
'polluter pays-principle'	0	0	0	1
'polluter pays' principle	7	46	218	35
'the polluter pays principle'	0	0	0	2
"polluter pays principle"	0	0	16	9
"polluter pays" principle	0	2	122	9
"the polluter pays principle"	0	0	1	0
"the polluter pays" principle	0	0	1	0
polluter pays	0	0	3	0
polluter pays principle	0	40	253	53
polluter principle	0	1	0	0
polluter should pay principle	0	1	0	0
polluter-pays	0	0	3	0
polluter-pays principle	0	0	15	1
Polluters pays principle	0	0	1	0
polluting pays principle	0	0	1	1
principle according to which the polluter should pay	0	0	2	0
principle of 'the polluter pays'	0	3	0	0
principle of "polluter pays"	0	0	2	0
principle of "the polluter pays"	0	1	0	0
principle of payment by the polluter	0	2	0	0
principle of polluter pays	0	0	1	0
principle of polluter-pays	0	0	1	0
principle of the polluter pays	0	2	1	0
principle relating to polluter pays	0	0	0	1
principle that the "polluter should pay"	0	0	1	0
principle that the polluter must pay	0	1	0	0
principle that the polluter pays	1	0	1	0
principle that the polluter should pay	1	10	19	10
principle that those responsible for the damage should pay	0	0	1	0
principle whereby the polluting party pays	0	0	1	0
principle which says that "the polluter pays"	0	0	1	0

Figure 1: English variants for the polluter pays principle.

The precautionary principle (Figure 2) also has a clear preferred variant: *precautionary principle*. In this case, terminological harmonization seems to be clearer, as the number of variants drops from 8 in the 1990s and 6 in the 2000s to 3 in the 2010s.
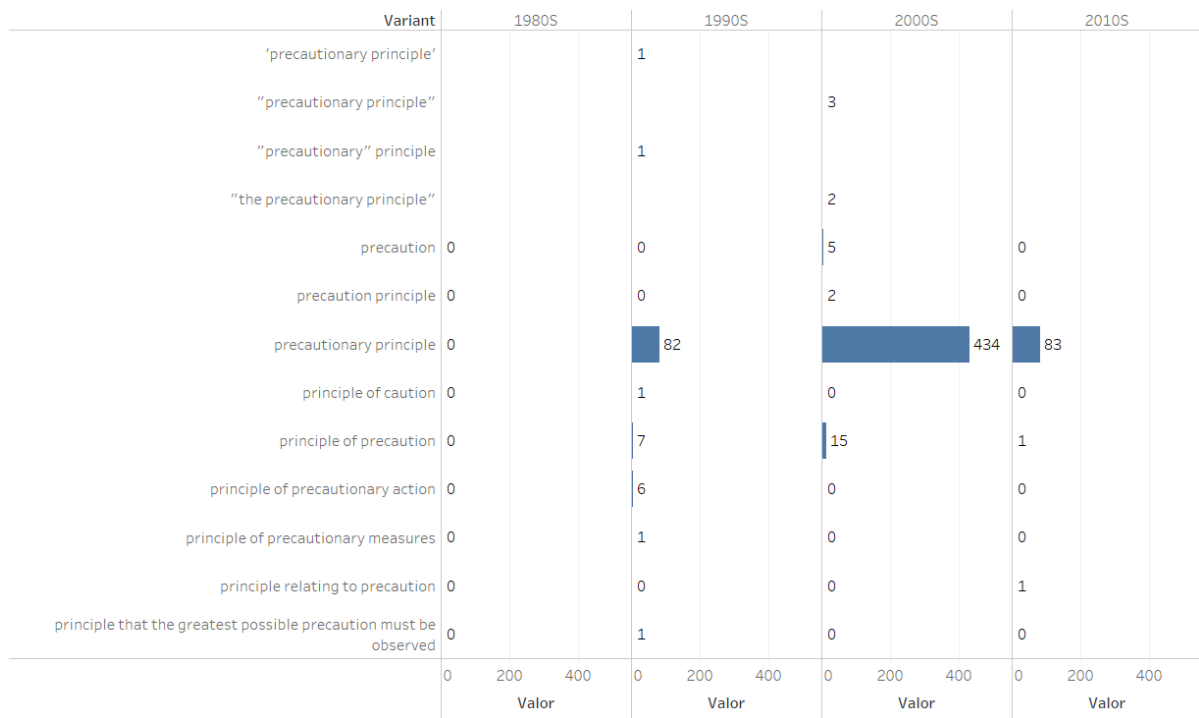
Figure 2: English variants for the precautionary principle.

Variant	1980S	1990S	2000S	2010S
'precautionary principle'		1		
"precautionary principle"			3	
"precautionary" principle		1		
"the precautionary principle"			2	
precaution	0	0	5	0
precaution principle	0	0	2	0
precautionary principle	0	82	434	83
principle of caution	0	1	0	0
principle of precaution	0	7	15	1
principle of precautionary action	0	6	0	0
principle of precautionary measures	0	1	0	0
principle relating to precaution	0	0	0	1
principle that the greatest possible precaution must be observed	0	1	0	0

The diachronic development of the prevention principle seems more chaotic according to our data (Figure 3). In the 1990s, three variants stand out: *principle that preventive action should be taken* (16), *principle of prevention* (14), and *principle of preventive action* (13). In the 2010s, however, the preferred variants are *principle that preventive action should be taken* (35), *principle of prevention* (22), *preventive principle* (19), and *prevention principle* (15). Then in the 2010s, only the *principle that preventive action should be taken* (13) perseveres, which may be the result of terminological harmonization, but may also be caused by the smaller size of the corpus and the lower number of results (26 in total).

The results show that the term formation process has only just started for this principle, and the same can be said for the rectification at source principle (Figure 4), as the preferred variant in all decades is *principle that environmental damage should [...] be rectified at source.* However, in the 2010s a new variant emerges, which seems slightly more lexicalized: *principle of rectification of pollution at source.*
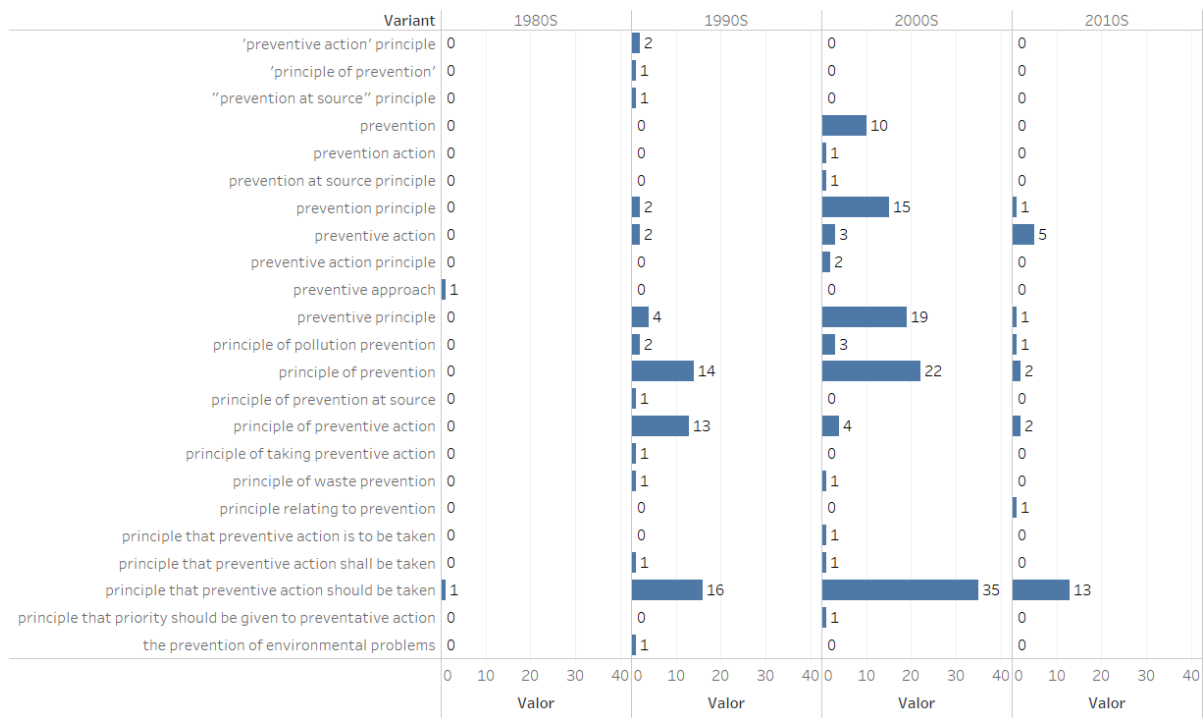
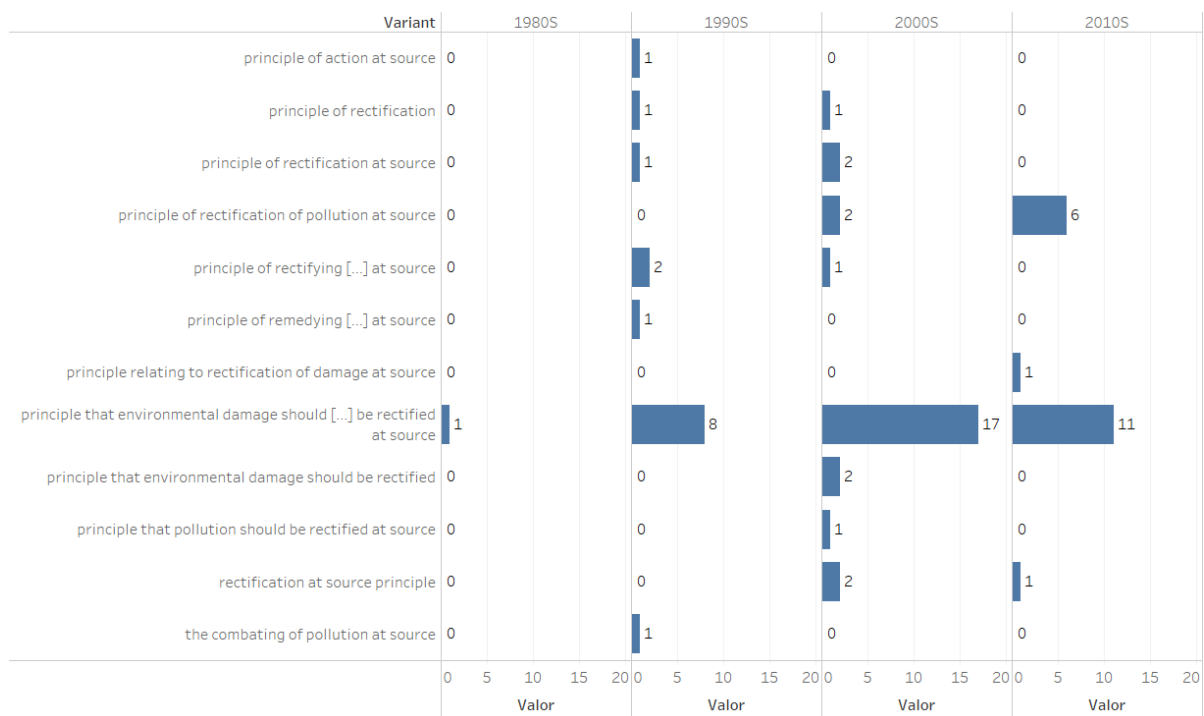**Figure 3:** English variants for the prevention principle.

Variant	1980S	1990S	2000S	2010S
'preventive action' principle	0	2	0	0
'principle of prevention'	0	1	0	0
"prevention at source" principle	0	1	0	0
prevention	0	0	10	0
prevention action	0	0	1	0
prevention at source principle	0	0	1	0
prevention principle	0	2	15	1
preventive action	0	2	3	5
preventive action principle	0	0	2	0
preventive approach	1	0	0	0
preventive principle	0	4	19	1
principle of pollution prevention	0	2	3	1
principle of prevention	0	14	22	2
principle of prevention at source	0	1	0	0
principle of preventive action	0	13	4	2
principle of taking preventive action	0	1	0	0
principle of waste prevention	0	1	1	0
principle relating to prevention	0	0	0	1
principle that preventive action is to be taken	0	0	1	0
principle that preventive action shall be taken	0	1	1	0
principle that preventive action should be taken	1	16	35	13
principle that priority should be given to preventative action	0	0	1	0
the prevention of environmental problems	0	1	0	0

Figure 4: English variants for the rectification at source principle.

Variant	1980S	1990S	2000S	2010S
principle of action at source	0	1	0	0
principle of rectification	0	1	1	0
principle of rectification at source	0	1	2	0
principle of rectification of pollution at source	0	0	2	6
principle of rectifying [...] at source	0	2	1	0
principle of remedying [...] at source	0	1	0	0
principle relating to rectification of damage at source	0	0	0	1
principle that environmental damage should [...] be rectified at source	1	8	17	11
principle that environmental damage should be rectified	0	0	2	0
principle that pollution should be rectified at source	0	0	1	0
rectification at source principle	0	0	2	1
the combating of pollution at source	0	1	0	0

The results show that there is a certain tendency towards terminological harmonization in the case of environmental law principles in English in the EUR-Lex2/2016 corpus, although the proliferation of term variants in all cases is quite astonishing. In addition, although there has been an effort for legal harmonization of environmental law in the EU, the term formation process of some of the principles is at its very beginning, even in English.

3.3. Diachronic variant analysis in Spanish

To study the diachronic evolution of Spanish variants and their possible terminological harmonization over time, we started out analyzing three common variants of the polluter pays principle in English, *polluter pays principle, 'polluter pays' principle*, and *"polluter pays" principle*, with the View translations in Parallel Concordance function. To do this, we used a simplified version of CQL 3:

[lemma="polluter"][]{0,5}[lemma="principle"] [tag!="N.*|JJ.*"] within <s/>

Sub-hits were again filtered out. Our analysis showed no correlation between specific English variants and their counterparts in the Spanish segments. For example, the English variants *"polluter pays" principle* and *polluter pays principle* both showed the same Spanish variants: *principio de "quien contamina paga", principio de que quien contamina paga, principio de quien contamina paga, principio "quien contamina paga", principio de "quien contamina, paga", principio de que "quien contamina, paga", principio de quien contamina, paga*, etc. The fact that the first English variant has quotation marks and the second does not, in no way predicts the use of quotation or other punctuation marks in the Spanish texts. There was also no clear preference for one equivalent for each English variant.

Therefore, we decided to accumulate the data and analyze term evolution of the Spanish variants separately. Figure 5 shows that the preferred Spanish variants for the polluter pays principle are *principio "quien contamina, paga", principio de "quien contamina, paga", principio de que "quien contamina, paga"*, and *principio de que quien contamina, paga* in the 1990s, 2000s, and 2010s. A slight tendency towards terminological harmonization can be seen as there seems to be a preference for three variants in the 2010s (*principio "quien contamina paga", principio de "quien contamina paga"*, and *principio de que quien contamina, paga*) as compared to five in the 2000s. The number of different variants grows from the 1990s (18) to the 2000s (23) and then decreases in the 2010s (14), which is also a sign of harmonization. However, the number of variants is again surprising.
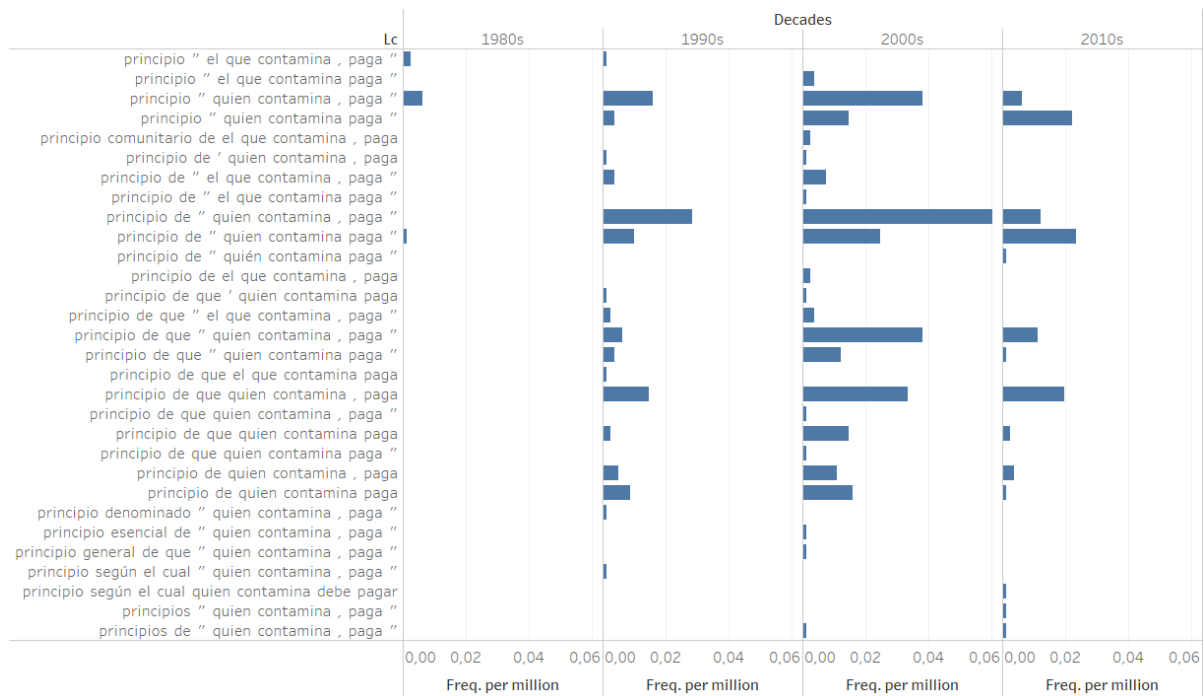

Figure 5: Spanish variants for the polluter pays principle.

One reason for the proliferation of Spanish variants in the case of the polluter pays principle may be that it includes a verbal form. In Spanish, this leads to many possible grammatical structures. Therefore, the harmonization process may take more time.

4. Conclusions and future research

In this study, we analyzed the diachronic evolution of term variants of environmental law principles in English and Spanish in the EUR-Lex2/2016 corpus. We assumed that legal harmonization would in time also lead to terminological harmonization. The results show that term variation is ubiquitous in the English as well as in the Spanish texts. Several reasons may explain this phenomenon. Firstly, it may be that not enough time has passed for harmonization to occur. The lack of lexicalization of the prevention principle and the rectification principle in English are certainly an indication in that sense. The fact that they are multiword terms, and in some cases even include verbs among their elements (for example the polluter pays principle), gives rise to term variation in Spanish as well [10]. Secondly, the institutional translation memories and termbases may not provide enough information on the preferred term variant in each case, especially when referring to the use of quotation and other punctuation marks. For example, IATE, EU's terminology management system, does not provide any information in that sense. Finally, individual translators may be under too tight time restrictions for them to make a rigorous selection of variants.

As ensuring terminological consistency is essential in international organizations for the sake of legal univocity and certainty, and thus for quality assurance [5], we believe that better term management practices must be insisted on. The term variants and equivalences provided by IATE, for example, do not provide enough information for sound variant selection. Many of the variants found in the present study are not present in the termbase and some of the variants it provides have not been found in our analysis. Figure 6, shows the IATE entry for the polluter pays principle. As main term it provides the hyphenated form *polluter-pays principle*, whereas that variant has not been the most used in any decade of our corpus, and the acronym. It also provides the Spanish equivalents "principio de responsabilidad económica del contaminador" and "principio del contaminador responsable", which were not encountered in our corpus.

Figure 6: IATE entry for *polluter-pays principle*.

For future research, we will carry out a detailed contrastive analysis of punctuation usage between English and Spanish to try to find an explanation of punctuation usage in our data. We will also relate our results with existing systematic classifications of term variation, such as the one proposed by [11]. Furthermore, we will look into the Spanish variants of the other environmental law principles to see if they behave differently and to see the influence of the level of lexicalization. Another line of work will be to relate variants to the text type in which they appear. Finally, we will create an additional corpus of EUR-Lex texts to study tendencies after 2016.

Acknowledgements

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] A. Kiss, International Environmental Law, 2nd. ed., Transnational Publishers, New York, NY, 2000.

[2] I. Przybojewska, Methods of Harmonisation in EU Climate and Energy Directives and their Impact on the Content and Interpretation of National Law, Rocznik Administracji Publicznej 4 (2018) 117–133. doi: 10.4467/24497800RAP.18.006.9221.

[3] Ł. Biel, Lost in the Eurofog: The Textual Fit of Translated Law, Peter Lang, Frankfurt am Main, 2014. doi: 10.3726/978-3-653-03986-3.

[4] D. Guzmán and F. Prieto Ramos, Assessing Legal Terminological Variation in Institutional Translation: The case of national court names in the human rights monitoring procedures of the United Nations, Translation and Translanguaging in Multilingual Contexts 7(2) (2021) 224–247. doi: 10.1075/ttmc.00067.guz.

[5] F. Prieto Ramos, International and Supranational Law in Translation: From Multilingual Lawmaking to Adjudication, The Translator 20(3) (2014) 313–331. doi: 10.1080/13556509.2014.904080.

[6] V. Baisa, J. Michelfeit, M. Medve, and M. Jakubíček, European Union Language Resources in Sketch Engine, in: Proceedings of the Tenth International Conference on Language Resources and Evaluation, LREC'16, ELRA, Portorož, Slovenia, 2016, pp. 2799–2803.

[7] A. Kilgarriff, V. Baisa, J. Bušta, M. Jakubícek, V. Kovár, J. Michelfeit, P. Rychlý, and V. Suchomel, The Sketch Engine: Ten Years On, Lexicography 1(1) (2014) 7–36.

[8] M. Cabezas-García and P. León-Araúz, Term and concept variation in climate change communication, The Translator 28(4) (2022) 429–449. doi: 10.1080/13556509.2023.2182168.

[9] L. Bowker and S. Hawkins, Variation in the organization of medical terms. Exploring some motivations for term choice, Terminology 12(1) (2006) 79110.

[10] P. León-Araúz, M. Cabezas-García, and A. Reimerink, Representing Multiword Term Variation in a Terminological Knowledge Base: a Corpus-Based Study, in: Proceedings of the 12th Conference on Language Resources and Evaluation, LREC 2020, ELRA, Marseille, France, 2020, pp. 2351–2360.

[11] P. Faber and P. León-Araúz, Specialized knowledge representation and the parameterization of context, Frontiers in Psychology 7(196) (2016) 1–20.

Online sources

EUR-Lex parallel corpus, https://www.sketchengine.eu/eur-lex-parallel-corpus/, available in Sketch Engine
IATE (Interactive Terminology for Europe), available at https://iate.europa.eu/home
Sketch engine, available at https://www.sketchengine.eu/