Understanding and translating the concept of DAMAGE in environmental texts

Useful writing assistance tools should facilitate specialized text encoding and translation. Accordingly, the new generation of terminological resources should include cognitive structures as well as syntagmatic information of various types, such as collocations, verb-noun combinations, complex noun terms and scientific phraseology (Faber 2012, 2015; L'Homme et al 2014, inter alia).

As is well known, language is much less compositional than one might initially think (L'Homme 2004, Williams 2005, Granger & Meunier 2008). This idea comes from a linguistic tradition (Firth 1961; Sinclair 1991; Hanks 2004; Halliday et al 2014) that states that selectional restrictions or "phraseological patterns" are crucial in order to describe words meaning and behavior (Hanks 2004, 2012).

Given the fact specialized texts are characterized by a high frequency of standardized lexico-grammatical structures, the study of phraseology in scientific writing is a topic that has attracted much interest in recent years (Tutin 2014, Gledhill 2000, Kübler & Pecman 2012).

In this sense, semantic categories are a useful way to describe syntagmatic patterns and disambiguate word senses. As reflected in resources such as WordNet, VerbNet, PropBank, FrameNet, etc., the different meanings of a verb can be explained based on semantic category restrictions (Hanks & Pustejovsky 2005: 73-74):

> *To grasp*
> 1. Seize hold of something: [Person] grasp [Physical Object]: *He grasped the handle of the door in one hand, and that of the spoon in the other.*
> 2. Understand: [Person] grasp [Abstract]: *She grasped what was happening.*

Following this perspective, this study focuses on the description of the semantic category of DAMAGE in the field of Environmental Science from a lexicographic perspective. Although much has been said about DAMAGE from a scientific point of view, its processes, agents, causes, and effects have still not been specified from a linguistic perspective for translation purposes.

In this research, an English corpus of 23 million words in English has been used to study DAMAGE. The results obtain show that this category pervades a wide range of Environmental Science subfields such as Coastal Engineering, Meteorology, Climatology Oceanography, Air Quality Management, Hydraulic Engineering, Environmental law, Water Treatment, etc).

In order to understand DAMAGE and its syntagmatic patterns, we have extracted its semantic participants (causes, consequences, attributes, affected entities, prototypical actions and events) as well as their lexicalizations and lexico-

grammar structures of this term both as a noun and a verb. The following table shows an excerpt of our results.

causes	*storm, tornado, hurricane, earthquake, tsunami, rain, wind, flood, erosion, climate change, fire, wave, spill, pollution, mudflow.*
consequences	*death, injury, flooding, life, casualties, disruption, toll, destruction, failure, disease, fatalities, cost, cancer, risk, pollution, mortality*
affected entities	*building, property, infrastructure, ecosystem, crop, vegetation, environment, plant, home, dune, health, life, community, coral reefs*
attributes	*structural, environmental, irreversible, extensive, severe, considerable, serious, ecological,*
actions & events	*cause, suffer, inflict, prevent, sustain, avoid, minimize, repair, assess, trample, reduce, estimate, limit, mitigate, expect, induce, incur, result, minimize*

It is generally assumed that a term can have different collocates depending on the features activated in different contexts (Sánchez Cárdenas & Buendía Castro 2012, San Martín & León-Araúz 2013). Consequently, we have analyzed the collocation patterns of DAMAGE in the different subdomains where it is activated. Thus, we have identified the different syntagmatic configurations of this concept, as shown in the following example:

Verb	Agent	Result	Patient	Semantic class structure
cause	*UV-A*	*damage*	*DNA*	[CHEMICAL ENTITY] [EVENT] [BIOLOGICAL STRUCTURE]
suffer	*hurricane*	*damage*	*building*	[NATURAL DISASTER] [EVENT] [ARTIFICIAL OBJECT]
inflict	*cyclone*	*damage*	*beach*	[NATURAL DISASTER] [EVENT] [NATURAL OBJECT]
prevent	*legislation*	Ø	*damage*	[TEXT][EVENT]
avoid	*antipollution technology*	*health*	*damage*	[PROCESS] [ATTRIBUTE] [EVENT]
reduce	*flood forecasting system*	*disaster mitigation*	*flood damage*	[PROCESS][EVENT][EVENT]

Additionally, the corpus shows that the noun *damage* can behave as the hypernym of specific types of damage such as *flooding, pollution, drought* or *biodiversity reduction*. In the second stage of our research, we extracted all the terms within this semantic class using two strategies. Firstly, we used the Sketch Engine[1] thesaurus tool to calculate the possible synonyms of *damage,* using distributional semantics. Secondly, knowledge patterns were used to extract hyponyms, using the sketch grammar developed by Leon et al (2016).

---

[1] [https://www.sketchengine.co.uk]

Both strategies were combined to extract all the members of this semantic class.

To translate a text on environmental damage translators must first understand the conceptual patterns underlying this concept. They must then find how these patterns are lexicalized in the target language. In this second phase, the use of corpora is extremely effective. Our research shows how the cognitive structures underlying this concept can be highlighted. Additionally, it describes its lexico-grammatical structures in English and their correspondences in Spanish. This procedure can be reproduced in the Translation classroom for didactic purposes.

References

Faber, P. (eds). (2012). *A cognitive linguistics view of terminology and specialized language* (Vol. 20). Walter de Gruyter.

Faber, P. (2015.) Frames as a framework for terminology. *Handbook of Terminology*, 1(14). ed. by Kockaert, H.J. & Steurs, F., 1:14-33. John Benjamins Publishing Company.

Firth, J. R. (1961). *Papers in Linguistics 1934-1951.* Oxford University Press.

Gledhill, Christopher J. (2000). *Collocations in science writing* (Vol. 22). Gunter Narr Verlag.

Granger, S., and Meunier, F. (eds). (2008). *Phraseology: An interdisciplinary perspective*. John Benjamins Publishing.

Halliday, M., Matthiessen, C. M., and Matthiessen, C. (2014). *An introduction to functional grammar*. Routledge.

Hanks, P., & Pustejovsky, J. (2005). A pattern dictionary for natural language processing. *Revue Française de linguistique appliquée*, *10*(2), 63-82.

Hanks, P. (2004). Corpus pattern analysis. In *Euralex Proceedings* (Vol. 1): 87-98.

Hanks, P. 2012. How people use words to make meanings: Semantic types meet valencies. *Input, Process and Product: Developments in Teaching and Language Corpora*, 54-69.

Kübler, N., and Pecman, M. (2012). *The ARTE bilingual LSP dictionary: From collocation to higher order phraseology*. na.

L'Homme, M. C., Robichaud, B., & Rüggeberg, C. S. (2014). Discovering frames in specialized domains. In *LREC* (pp. 1364-1371).

L'Homme, M.C. (2004). A Lexico-semantic Approach to the Structuring of Terminology. In *Proceedings of Computerm*: 7-14.

León-Araúz, P., San Martín, A. & Faber, P. (2016) Pattern-based Word Sketches for the Extraction of Semantic Relations. In *Proceedings of the 5th International Workshop on Computational Terminology (Computerm2016)*, pages 73-82. Osaka, Japan: COLING 2016

San Martín, A., & Araúz, P. L. (2013). Flexible Terminological Definitions and Conceptual Frames. In *VDOS+ DO@ ICBO*.

Sánchez Cárdenas, B. & Buendía Castro, M. (2012) Inclusion of Verbal Syntagmatic Patterns in Specialized Dictionaries: The Case of EcoLexicon. In *Proceedings of the 15th EURALEX International Congress*. Ruth Vatvedt Fjeld and Julie Matilde Torjusen, pages 554-562. Oslo: EURALEX.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Tutin, A. (2014). *L'écrit scientifique: du lexique au discours*. F. Grossmann (Ed.). Presses universitaires de Rennes.

Williams, G. (2005). English Collocation Studies: The OSTI report. *International Journal of Lexicography*, 18(3): 391-393.