Culture-bound Representation in EcoLexicon (Short Paper)

Pilar León-Araúz[1], Melania Cabezas-García[1], Pamela Faber[1] and Arianne Reimerink[1]

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

Abstract

Culture permeates all aspects of life, and even influences our perception of the world. Culture is also reflected in specialized communication. Since terms and their meanings are largely dependent on culture, the issue is how to account for this cultural component in terminological resources. This paper explains how this can be achieved in EcoLexicon (http://ecolexicon.ugr.es), a multilingual terminological knowledge base on environmental science. This involves specifying a set of cultural profiles or frames 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 entails adding a cultural component to all modules (definitions, semantic networks, terms, phraseology, and multimodal resources). Therefore, new cultural parameters must be added to culturally expand or restrict contextual representations. These parameters would refer to geographic location, domain, and degree of specialization. Cultural representation in EcoLexicon will take the form of flexible definitions, dynamic semantic networks, contextually-enriched term entries, and images selected and displayed based on contextual criteria.

Keywords

Culture-bound terminology, EcoLexicon, definitions, semantic networks, images

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 [1]. Not surprisingly, culture is also reflected in specialized language and terminology. Recently, the cultural facet of Terminology or Culture-Bound Terminology [2] has been highlighted by Temmerman and Van Campenhoudt [3], Faber and Medina-Rull [4], and Diki Kidiri [5]. 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 [6]. Since terms and their meanings are culturally motivated, the issue is how to represent this cultural dimension in terminological knowledge bases. This paper explains how this can be achieved in EcoLexicon (http://ecolexicon.ugr.es), a multilingual environmental knowledge base.

2. Cultural representation in EcoLexicon

EcoLexicon represents the conceptual structure of the specialized domain of the Environment in the form of a visual thesaurus in which environmental concepts are configured in semantic networks.

2nd International Conference on "Multilingual digital terminology today. Design, representation formats and management systems" (MDTT) 2023, June 29–30, 2023, Lisbon, Portugal

EMAIL: pleon@ugr.es (P. León-Araúz); melaniacabezas@ugr.es (M. Cabezas-García); pfaber@ugr.es (P. Faber); arianne@ugr.es (A. Reimerink)

ORCID: 0000-0002-8520-2749 (P. León-Araúz); 0000-0002-8622-1036 (M. Cabezas-García); 0000-0003-0581-0005 (P. Faber); 0000-0002-7264-4580 (A. Reimerink)

The objective is to convert EcoLexicon into an inclusive resource sensitive to cultural variation. This involves specifying a set of cultural profiles or frames 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, semantic networks, terms, phraseology, and multimodal resources). Since concepts in EcoLexicon can already be recontextualized by domain (e.g. Geology, Coastal Engineering, etc., which reflect sociocultural usage), new cultural parameters must be added to culturally expand or restrict contextual representations. These parameters would refer to geographic location (e.g. United States vs. United Kingdom, Spain vs. Spanish-speaking countries in the Caribbean), and degree of specialization (expert, semi-expert, and non-expert). However, since not all domains are equally represented in EcoLexicon, we will focus on those which are particularly subject to cultural variation (e.g. Biodiversity). The integration of these parameters in each module requires different strategies as well as an in-depth reflection on terminological equivalence, denominative variation, and culturally specific concepts, terms and phrasemes. The data represented will be mined from English and Spanish subcorpora of the EcoLexicon corpus [7] that are composed of texts representing each cultural parameter.

2.1. Definitions and semantic networks

Cultural variation is usually reflected in multidimensional concepts, whose relational behavior changes based on contextual parameters. For example, the concept WATER has an active role in Geology (it causes erosion, reshapes the terrestrial landscape, etc.), while WATER in the Water Treatment domain is a patient that receives actions (purification, filtering, etc.). This multidimensionality will be reflected in the form of flexible definitions [8, 9] and dynamic semantic networks in EcoLexicon.

Each concept in EcoLexicon is described by a definition, which is a natural language explanation of its location in the conceptual structure of a domain. A definition not only specifies the properties of a concept, but also links it to others [10], thus providing a blueprint for the concept and its semantic network. The recontextualization of definitions and semantic networks involves creating contextual constraints [11, 8]. This has already been done by applying domain-specific restrictions to certain general concepts such as WATER and SAND [12, 13]. Table 1 shows how the definitions of SAND can be adapted based on the domain where SAND is activated. For example, while SAND is a sediment in Geology, it is regarded as a soil component in Soil Sciences. Specialized domains are thus indicative of sociocultural usage/conceptualization since groups of people use language and relate concepts differently.

Table 1
Definitions of SAND in different domains

SAND			
General environmental definition	Unconsolidated **mineral material** consisting mainly of fragments of quartz ranging in size of 0.05-2 mm.	*GEOLOGY* definition	**Sediment** consisting mainly of fragments of quartz ranging in size of 0.05-2 mm that is part of the soil and can be found in great quantities in beaches, river beds, the seabed, and deserts.
SOIL SCIENCES definition	Unconsolidated **soil component** consisting mainly of fragments of quartz ranging in size of 0.05-2 mm that are the result of weathering and erosion.	*CIVIL ENGINEERING* definition	**Natural aggregate** consisting mainly of fragments of quartz ranging in size of 0.05-2 mm that is a component of diverse construction material such as concrete and mortar.

Accordingly, cultural recontextualization depends on a set of cultural parameters, based on geographic location, historical time period, sociocultural usage, etc. which restrict the definitions to a certain cultural context. The result will be a set of culturally-adapted, flexible definitions. For example, the LIONFISH is present in different geographic locations, such as the Indo-Pacific and the Caribbean Sea. A general description of LIONFISH based on its family, origin, and physical characteristics is not the same as a Caribbean-specific definition, where it is a predator and invasive species that has a negative impact on native species and their habitat.

Adapting definitions would necessarily entail expanding the set of semantic relations in the knowledge base and providing dynamic semantic networks, which only display the relations encoded in a particular context. Thus, users could choose a sociocultural restriction of interest (e.g. a particular domain, geographic location, etc.) and only relations activated in that context would be displayed. For instance, WATER *causes* EROSION in the Geology domain (Figure 1) rather than in the Water Treatment domain, where other relations are encoded, such as PURIFICATION *affects* WATER. For that purpose, contextual restrictions will be applied to conceptual propositions in the database.

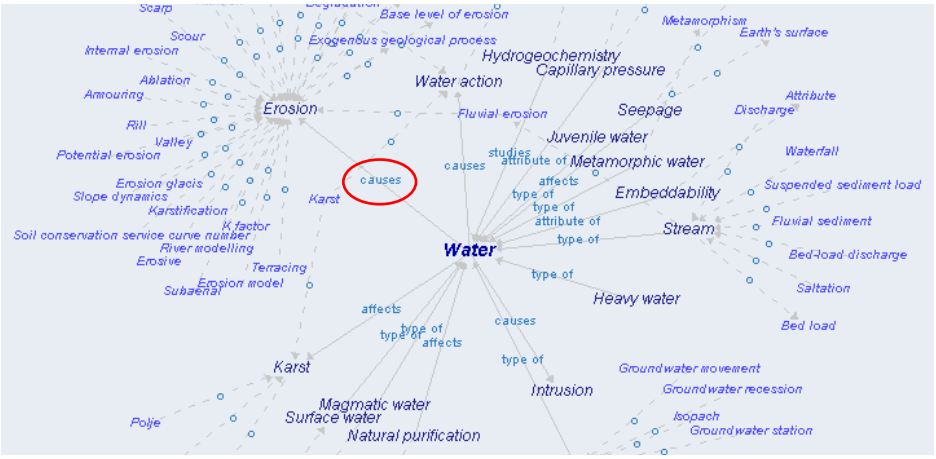

Figure 1: WATER in the Geology domain

An example of restrictions in semantic networks for a concept that behaves differently according to its geographical location is WETLAND. In Figure 2, the network to the left shows the general network for WETLAND, whereas the network to the right is restricted for the Caribbean, with MARSH and SWAMP as prototypical wetlands for the area, and SEAGRASS BED, which is only there considered a wetland.

Figure 2: Non-restricted semantic network (left) and network restricted for the Caribbean (right) for WETLAND

2.2. Terms and phraseology

Terms also reflect culture in a variety of ways. For example, certain terms metaphorically refer to a familiar entity in a culture (e.g. *cookie-cutter shark*). The interest in such terms lies in their correspondences in other languages. For example, in Spanish, the *cookie-cutter shark* is a *tiburón cigarro* (lit. *cigarette shark*) because of its long thin shape and dark brown band around its gills, which makes it resemble a cigar. In contrast, the English term focuses on the perfect circularity of its bite and the process involved when it attacks its prey, which in English-speaking countries is a phase of cookie making.

Terms can also reflect the sociocultural level, geographic location or historical period of the speaker. At the very least, representing this type of information involves a set of pragmatic fields that define the type of denominative variation. Pragmatic fields will include the following: i) Use_Geographical (for terms that are used in specific geographic locations); ii) Use_Register (to show variation based on the degree of specialization); iii) Use_Context (information related to monolingual contexts different from register); and iv) Use_Translation_Context (usage information found in parallel corpora, such as the preference for a certain translation). In Ecolexicon, these fields will be visible in the information for each term (Figure 3). Furthermore, variants will also be displayed in a contrastive view, which compares them based on their frequency, and formal, cognitive and diachronic changes (Figure 4) [14].

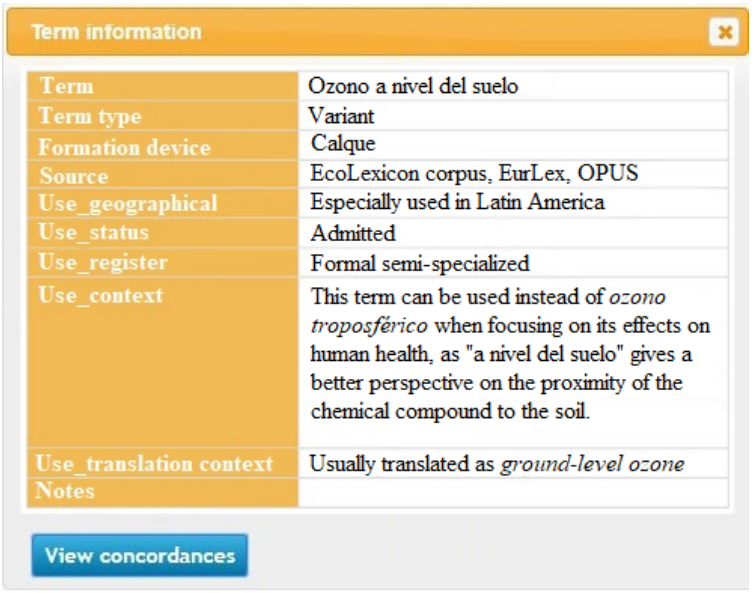

Figure 3: Term information for *ozono a nivel del suelo* (*ground-level ozone*)

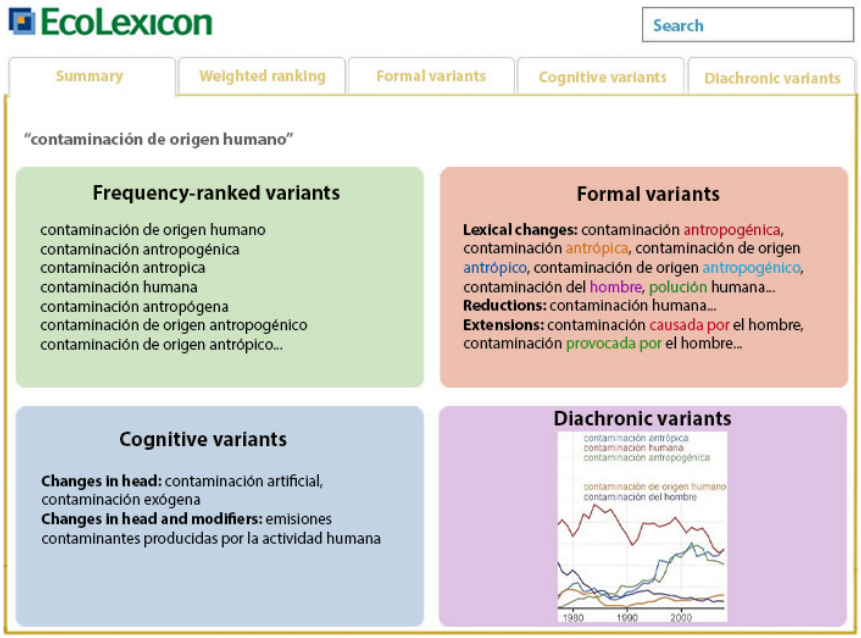

Figure 4: Contrastive view for *contaminación de origen humano* (*anthropogenic pollution*)

Moreover, in EcoLexicon, phraseological units are both verb collocations (e.g. *CFCs destroy ozone*) and multiword terms (e.g. *ozone depletion potential*), which can exhibit a varying degree of compositionality. Multiword terms are included in term entries, which will be enhanced with pragmatic information (see Figures 3 and 4). Verb collocations are also included in the terms that they include and will be annotated with sociocultural information. For instance, *CFCs* ***deplete*** *ozone* is a phraseological unit of a formal register, whereas *CFCs* ***destroy*** *ozone* is more informal.

2.3. Multimodal resources

In EcoLexicon, concept entries also contain multimodal information for the sake of knowledge acquisition because the combination of textual and visual material improves understanding [15]. With a view to transforming EcoLexicon into a culture-sensitive resource, the selection criteria for multimodal information, such as images, will also be based on the cultural context [6].

To this end, images will be tagged and displayed with one (or several) of the following contextual parameters: i) geographic location; ii) level of specialization; and iii) domain. Figure 5 shows three images for CORAL REEF differentiated by geographical location (the Caribbean [left] and Scandinavia [middle]) and by level of specialization (right), since textual features and other graphical information (eg. map) facilitate additional knowledge to lay users. Figure 6, on the other hand, shows an image for WATER specific for the Chemistry domain (left) and another for the Geology domain (right).

Figure 5: Image for CORAL REEF in the Caribbean (left) and in Scandinavia (right)

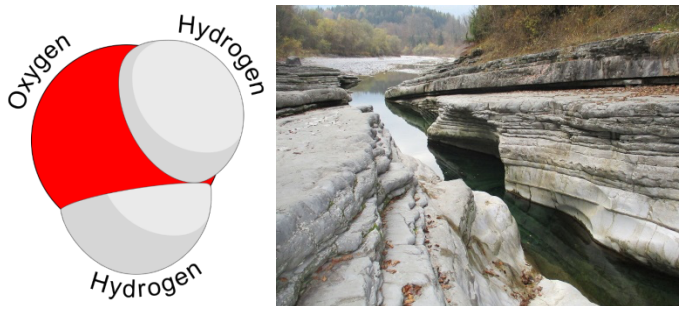

Figure 6: Image for WATER in the Chemistry domain (left) and Geology domain (right)

3. Conclusions

Culture plays a central role in communication. It is also present in specialized discourse, in the form of variation based on geographic location, specialized domain, and degree of specialization. This paper analyzes how EcoLexicon, a multilingual terminological knowledge base, can be transformed into a culture-sensitive resource. To this end, flexible definitions, dynamic semantic networks, contextually-enriched term entries, and images selected and displayed based on contextual criteria are proposed.

Given the socio-economic impact of environmental issues, the benefits of the descriptive approach in EcoLexicon should be highlighted, with a view to disseminating knowledge and raising awareness about environmental issues. The cultural design of EcoLexicon will describe a new vision on the environment, which acknowledges different cultural identities and highlights environmental problems around the world.

4. Acknowledgements

This research was carried out as part of the projects PID2020-118369GBI00 and A-HUM-600-UGR20, funded by the Spanish Ministry of Science and Innovation and the Regional Government of Andalusia.

5. References

[1] S. J., C. R. Sears, P. M. Pexman, Cultural Influences on Categorization Processes, Journal of Cross-Cultural Psychology 36, 6 (2005) 662-688. https://doi.org/10.1177/0022022105280509

[2] M. Diki-Kidiri: Le vocabulaire scientifique dans les langues africaines. Pour une approche culturelle de la terminologie. Paris, Karthala, 2008.

[3] R. Temmerman, M. Van Campenhoudt: Dynamics and Terminology. Amsterdam, Philadelphia, John Benjamins, 2014.

[4] P. Faber, L. Medina-Rull, Written in the Wind: Cultural Variation in Terminology. In: M. Gryviel (Ed.), Cognitive Approaches to Specialist Languages, Cambridge Scholars, Newcastle-upon-Tyne, 2017, pp. 419–442.

[5] M. Diki-Kidiri: Cultural Terminology. An introduction to theory and method, in: P. Faber, M. C. L'Homme (Eds.), Theoretical Perspectives on Terminology: Explaining terms, concepts and specialized knowledge, John Benjamins, Amsterdam, Philadelphia, 2022, pp. 197–216. https://doi.org/10.1075/tlrp.23.09dik

[6] M. Cabezas-García, A. Reimerink, Cultural Context and Multimodal Knowledge Representation: Seeing the Forest for the Trees, Frontiers in Psychology 13 (2022) 1–16. doi:https://doi.org/10.3389/fpsyg.2022.824932.

[7] P. León-Araúz, P., A. San Martín, A. Reimerink, The EcoLexicon English Corpus as an open corpus in Sketch Engine, in: J. Čibej, V. Gorjanc, I. Kosem, S. Krek (Eds.), Proceedings of the 18th EURALEX International Congress, Euralex, Ljubljana, 2018, pp. 893–901.

[8] A. San Martín, La representación de la variación contextual mediante definiciones terminológicas flexibles, Ph.D. tesis, Universidad de Granada, Granada, 2016.
[9] A, San Martín. A Flexible Approach to Terminological Definitions: Representing Thematic Variation. International Journal of Lexicography 35 (1), 53-74, 2022.
[10] B. E. Antia: Terminology and Language Planning. An alternative framework of practice and discourse. Amsterdam, Philadelphia, John Benjamins, 2000.
[11] A. San Martín, P. León Araúz, Flexible Terminological Definitions and Conceptual Frames, in: S. Seppälä, A. Ruttenberg (Eds.), Proceedings of the International Workshop on Definitions in Ontologies (DO 2013), Concordia University, Montreal.
[12] P. León Araúz, A. Reimerink, A. García Aragón, Dynamism and context in specialized knowledge, Terminology, 19.1 (2013): 31–61. doi:10.1075/term.19.1.02leo.
[13] P. León-Araúz, Term and concept variation in specialized knowledge dynamics, in: P. Drouin, A. Francœur, J. Humbley, A. Picton (Eds.), Multiple Perspectives on Terminological Variation, volume 18 of Terminology and Lexicography Research and Practice, John Benjamins, Amsterdam, Philadelphia, 2017, pp. 213–258. doi:10.1075/tlrp.18.09leo.
[14] P. León-Araúz, M. Cabezas-García, 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, 2020, pp.2351–2360.
[15] M. P. Cook, Visual representations in science education: the influence of prior knowledge and cognitive load theory on instructional design principles, Science Education 90 (2006) 1073–1091. doi: 10.1002/sce.20164