EcoLexicon

Pamela Faber, Miriam Buendía Castro
University of Granada
pfaber@ugr.es, mbuendia@ugr.es

Abstract

EcoLexicon (http://ecolexicon.ugr.es) is an environmental knowledge base which is based on the premises of Frame-based Terminology (FBT) (Faber 2009, 2011, 2012). 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. The various terminological designations for a concept are offered in six languages: Spanish, English, German, French, Russian, and Greek. So far, EcoLexicon contains 3,527 concepts and 18,617 terms. It provides conceptual, linguistic, and administrative information for each entry, and we are now beginning to include phraseological information for terms as well (Buendía 2013; Buendía & Sánchez 2012, Sánchez & Buendía 2012). EcoLexicon is designed to meet the needs of different user types, such as a student of science wishing to acquire specialized knowledge about a certain concept, a translator seeking translation correspondences in a language, or a specialist interested in text production.

Keywords: knowledge base; specialized language; environment

1   Introduction

EcoLexicon (http://ecolexicon.ugr.es) is a visual online thesaurus of environmental science, which currently contains 3,527 concepts and 18,617 terms in English, Spanish, German, French, Russian and Modern Greek. It provides conceptual, linguistic, and administrative information for each entry. This information as well as the corpus of specialized texts is stored in a private database, which allows members of EcoLexicon to add, eliminate, and/or modify conceptual and terminological information. In EcoLexicon it is assumed that up to a certain level, its potential users are familiar with scientific language and its usage in English or Spanish at least, since these are the interface languages. Potential users should thus possess a good command of any of the six languages in the knowledge base, as well as a minimum of scientific knowledge (López, Buendía & García 2012: 62).

EcoLexicon is based on the premises of Frame-based Terminology (FBT) (Faber 2009, 2011, 2012), a cognitive approach to Terminology. The FBT approach to Terminology applies the notion of frame, defined as a schematization of experience (a knowledge structure), which is represented at the conceptual level and relates elements and entities associated with a particular culturally embedded scene, situation or event from human experience.

The public version of EcoLexicon is freely available online. The new version of EcoLexicon includes the following new features, which are an improvement over the previous version.

- It is compatible with all modern browsers and does not require Java.
- It has more interactive and configurable maps, which allow the users to do the following:
  - Change the scale of the map.
  - Select the relations to be represented.
  - Eliminate nodes in order to make other nodes more prominent.
  - Adjust the position of nodes.
  - Go backwards or forwards.
  - Establish a direct link to a concept or term.
  - Look up a concept or term on Google or Google Images.
- It has new representation modes:
  - Hierarchical tree structure
  - Map of the shortest path between concepts
- It allows user registration to do the following:
  - Personalize results.
  - Store user preferences between sessions.

2 The micro and macrostructure of entries in EcoLexicon

Figure 1 displays the entry for *alluvial fan* in EcoLexicon. As shown in Figure 1, when EcoLexicon is opened, three zones appear:

(1) The top bar that allows access to different functionalities, such as the term/concept search or changing the language of the interface. It also permits to personalize the search by means of the contextual domains generally associated with a specialized knowledge field (e.g. Geology, Coastal Engineering, Environmental Law, etc.). This allows users to focus on the knowledge area or specific domain and to eliminate irrelevant information. It also allows users to create an account. This permits the storage of their preferences and options, independently of the computer used to access the database.

(2) The central area that includes a dynamic network that displays the search concept/term and links it to all related items in terms of a closed inventory of conceptual relations. When users click on any of the terms or concepts in the map, the network rearranges itself. In this new map, the term/concept that was clicked on is at the center and is connected with all of the entities directly related to it. As shown (Figure 1), in the lower left corner of the map, there is a text box with captions that allow users to identify the three categories of conceptual relations in EcoLexicon: (i) hyponymic (generic-specific) relations; (ii) meronymic (part-whole) relations; (iii) non-hierarchical relations. The conceptual relations used in EcoLexicon include a set of 17 hierarchical (hypernymic

and meronymic) and non-hierarchical relations, some of which are domain-specific. The 17 relations included are the following: AFFECTS, ATTRIBUTE_OF, CAUSES, COMPOSED_OF (MATERIAL), OPPOSITE_OF, DELIMITED_BY, STUDIES, PHASE_OF, MEASURES, PART_OF, REPRESENTS, RESULT_OF, EFFECTED_WITH, HAS_FUNCTION, TAKES_PLACE_IN, TYPE_OF, LOCATED_AT. In addition, a series of tabs are included in this central area that allow the user to access: (i) the history of concepts/terms visited; (ii) the results of the most recent search, which allows users to go backwards and forwards on this list, in the same way as on web browsers; (iii) the shortest path between two concepts; (iv) concordances for a term

Figure 1: Entry of *alluvial fan* in EcoLexicon.

(3) The side bar (Figure 2) that shows a wide range of information about the search item:
- The terminographic definition of the concept, which encodes the most basic conceptual relations in the category template;
- The associated terms designating the concept in the six languages contained in EcoLexicon. The list is ordered by language and term type (main entry term, synonym, variant, acronym, etc.). At the left of each term, there is a small flag representing the country where its language is spoken. By clicking on each term, a new window is displayed which provides users with the following linguistic information: (i) morphosyntactic information regarding grammatical category (noun, verb, adjective or adverb); term type (main term, synonym, geographical variant, and acronym); gender (masculine, feminine, and neuter); (ii) contexts of use; (iii) phraseology.
- The associated resources, (images, documents, URLs, audiovisual material, etc.) associated with each concept/term. Regarding images, depending on the type of knowledge activated, three different kind of images are distinguished, namely, iconic, abstract, or dynamic images (Prieto, 2009; Prieto &

López 2009). The complete list of the resources for each concept is shown in this box. Users can easily identify the type of resource by means of the icons beside each of the resource listed. In order to access more information regarding the resource (title, description, source, etc.) users can place the cursor on the resource and a new window will open with all this information.

- Conceptual categories. Each concept in EcoLexicon is associated with one or more conceptual categories, which are shown as a list. If users click on one of these categories, this opens a window with a list of all the concepts included. Furthermore, this box includes a *Category hierarchy* icon, which, when clicked on, shows the concepts in a hierarchical format in which nodes can expand or retract. If one of the categories in the hierarchy is clicked on, a window appears with all the concepts associated with that category.
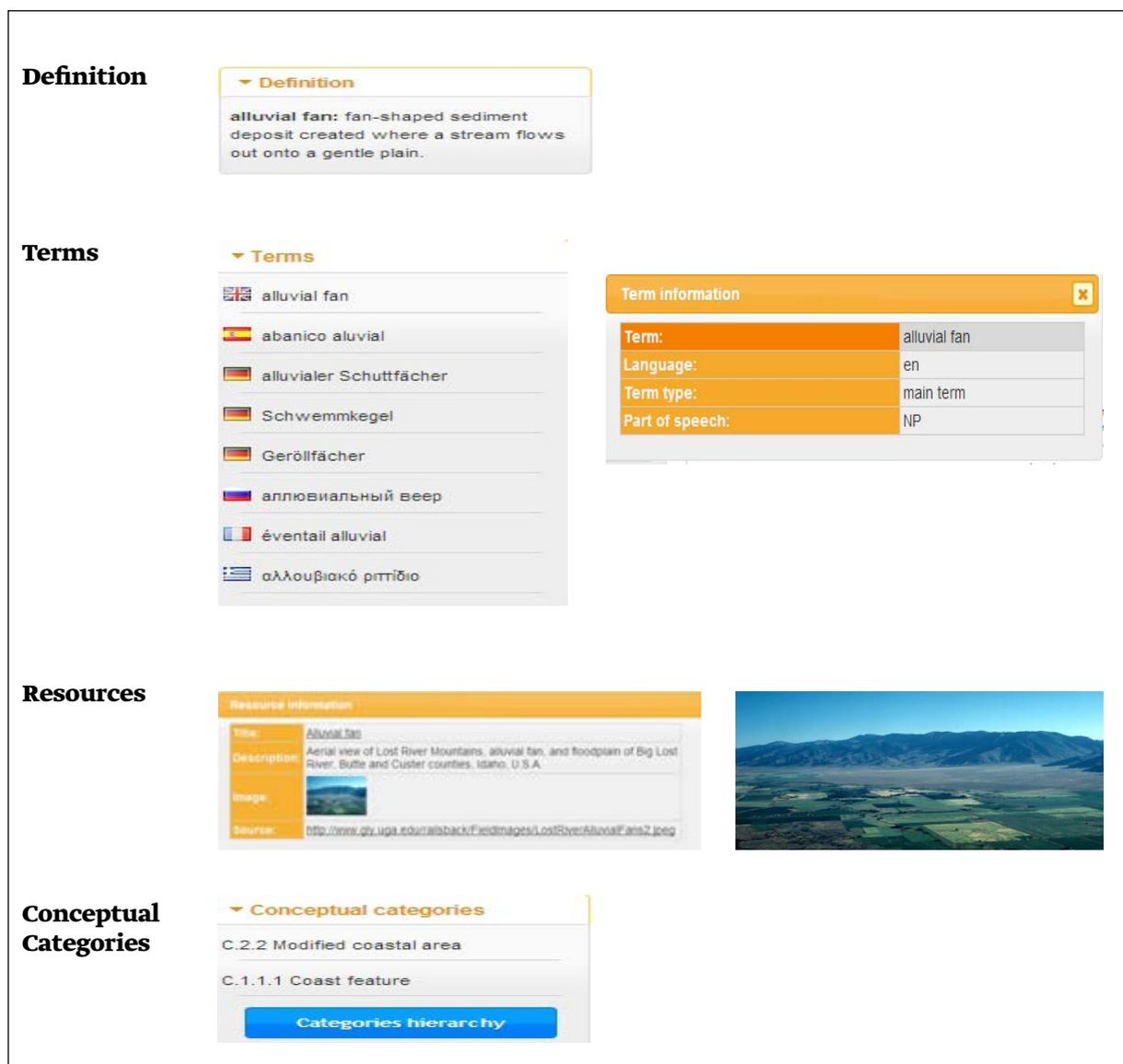
Figure 2: Extract of the side bar associated for each search item in EcoLexicon.

Figure 3: Extract of the phraseology box associated with *hurricane* in EcoLexicon.

- Phraseology. If the central element in the map is a term, this box shows a list of verbs most commonly used with the term. Phraseological information can also be accessed via the term entry in certain concepts. Verb collocations are classified and described according to meaning. For this reason, they were primarily classified in terms of their lexical domain (i.e. the nuclear meaning), and subsequently in terms of the frame activated within each lexical domain (i.e. meaning dimension). Once the lexical domain and frames are stated, the verbs are specified. By clicking on the verbs, the user has access to the usage sentences for the verb in question, as well as a note section with information about meaning restrictions.

3   Conclusion

As shown, Frame-based Terminology provides a full account of the information necessary to describe a specialized knowledge unit in a terminological entry. The practical application of FBT, the environmental knowledge base EcoLexicon, provides conceptual, linguistic, and administrative information for each specialized knowledge unit. To enhance knowledge acquisition,

conceptual information in EcoLexicon is stored and represented in different ways. Specialized environmental knowledge is represented by means of conceptual networks codified in terms of conceptual propositions in the form of a triplet (concept relation concept). The conceptual relations used in EcoLexicon include a set of 17 hierarchical (hypernymic and meronymic) and non-hierarchical relations. Conceptual information is also shown in the form of natural language definitions in English and Spanish, which are based on the conceptual propositions established by the concept to be defined. Additionally, domain-specific knowledge is also presented in the form of images, documents and videos, which complement the previously entered conceptual information. Linguistic information is now being enhanced by the introduction of verb phraseological information regarding each term. In a near future, we hope to be able to provide a complete phraseological description of every term.

4   References

Buendía Castro, M. (2013). Phraseology in Specialized Language and its Representation in Environmental Knowledge Resources. PhD Thesis, University of Granada, Spain.

Buendía Castro, M., & Sánchez Cárdenas, B. (2012). "Linguistic knowledge for specialized text production". In N. Calzolari, K.Choukri, T. Declerck,, M. Uğur Doğan, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis (eds.) Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12). Istanbul: ELRA, pp. 622-626.

Faber, Pamela. (2009). The Cognitive Shift in Terminology and Specialized Translation. *MonTI. Monografías De Traducción e Interpretación* 1(1), pp. 107–134.

Faber, Pamela. (2011). The Dynamics of Specialized Knowledge Representation: Simulational Reconstruction or the Perception-action Interface. *Terminology* 17(1), pp. 9–29.

Faber, Pamela (ed.). (2012). A Cognitive Linguistics View of Terminology and Specialized Language. Berlin, Boston: Mouton de Gruyter.

López Rodríguez, C.I., Buendía Castro, M. & García Aragón, A. (2012). User Needs to the Test: Evaluating a Terminological Knowledge Base on the Environment by Trainee Translators. *Jostrans. The Journal of Specialized Translation* (18), pp. 57–76. Accessed at: http://www.jostrans.org/issue18/art_lopez.pdf. [20/12/2012].

Prieto Velasco, J.A. (2009). Traducción e Imagen: La Información Visual En Textos Especializados. Granada: Tragacanto.

Prieto Velasco, J.A. & López Rodríguez, C.I. (2009). Managing graphic information in terminological knowledge bases, *Terminology* 15(2), pp. 179-213.

Sánchez Cárdenas, B., & Buendía Castro, M. (2012). Inclusion of Verbal Syntagmatic Patterns in Specialized Dictionaries: The Case of EcoLexicon. In R. Vatvedt Fjeld, J. Matilde Torjusen (eds.) Proceedings of the 15th EURALEX International Congress. Oslo: EURALEX, pp. 554–562.

Acknowledgements

608