

Complex Nominals in Terminology: Formation, Translation, and Representation

Melania Cabezas-García

University of Granada, SPAIN

melaniacabezas@ugr.es

ABSTRACT

This study has focused on English and Spanish complex nominals (CNs) pertaining to wind power terminology. Such multiword terms are very frequent in general and specialized discourse, but they are a specific feature of the latter since knowledge is usually condensed in these structures. Although CNs have been widely studied, there are various aspects that require further research. These include the formation of complex structures in specialized language (e.g. *wind power forecast error*), their modification by parts of speech different from the noun (e.g. *rotating magnetic field*), their presence in or translation into other languages, and their description in linguistic resources.

With a view to analyzing CN behavior in the wind power terminology, our specific objectives were the following: (i) to study CNs and address their characteristics in specialized discourse; (ii) to analyze CNs semantically; (iii) to explore CN formation from a semantic perspective; (iv) to investigate CN translation in English and Spanish; and (v) to design a systematic proposal for CN representation in the terminological knowledge base EcoLexicon. Assuming the premises of Frame-Based Terminology (Faber et al. 2005, 2006; Faber 2012), we compiled an English-Spanish comparable corpus on wind power, which was used for term extraction, semantic analysis, and the identification of terminological correspondences.

Our study revealed that the clusters of concepts in CNs can be accessed and specified by means of semantic categories, roles and relations. These also allowed us to capture regularities in CN formation, which can be further examined by a slot-filling mechanism activated by the CN head.

This triggers the so-called *microcontext*. CN translation was also found to be challenging because of the different patterns of term formation and the lack of systematicity regarding the description of CNs in terminographic resources. Along these lines, corpus techniques and guidelines for the specification of equivalences (based on an in-depth contrastive analysis) are essential. As for CN representation, it should encompass the different characteristics of these multiword terms (i.e. their morphosyntactic and semantic formation, their structure, their variants and equivalents, their position in the conceptual system, etc.). This wealth of information, together with a conceptual organization and different access paths to data, can provide a full account of specialized CNs and improve user experience. The practical application of this research is a CN section that has been devised for the phraseological module of EcoLexicon, a terminological knowledge base on the environment.

KEYWORDS: complex nominal, corpus linguistics, microcontext, multiword term, terminology, translation

Completion of Thesis

Place: University of Granada, Spain

Year: December 2019

Supervisors: Prof. Dr. Pamela Faber and Prof. Dr. Pilar León-Araúz

Original Language: Spanish

Los Compuestos Nominales en Terminología: Formación, Traducción y Representación

Melania Cabezas-García

University of Granada, SPAIN

melaniacabezas@ugr.es

RESUMEN

Este estudio se centra en los compuestos nominales (CN) del inglés y el español pertenecientes a la terminología de la energía eólica. Se trata de términos compuestos que son muy frecuentes en el discurso general y en la lengua especializada, pero que son especialmente característicos en esta última, como transmisores del conocimiento especializado. Aunque estos términos han recibido bastante atención, existen varios aspectos por explorar. Entre ellos, cabe destacar la formación de estructuras complejas en el lenguaje especializado, su modificación por categorías gramaticales distintas del sustantivo, su presencia en otras lenguas, o su traducción y descripción en los recursos lingüísticos.

Con el fin de analizar el comportamiento de los CN en la terminología eólica, nuestros objetivos específicos fueron los siguientes: (i) estudiar los CN y abordar sus características en el discurso especializado; (ii) analizar semánticamente los CN; (iii) explorar la formación de CN desde una perspectiva semántica; (iv) investigar la traducción de CN en inglés y español; y (v) diseñar una propuesta sistemática de representación de CN en la base de conocimiento terminológica EcoLexicon. Asumiendo las premisas de la Terminología Basada en Marcos (Faber et al. 2005, 2006; Faber 2012), compilamos un corpus comparable inglés-español sobre energía eólica, que se utilizó para la extracción de términos, el análisis semántico y la identificación de correspondencias terminológicas.

Nuestro estudio reveló que es posible acceder a los grupos de conceptos presentes en los CN empleando categorías, roles y relaciones semánticas. Estas también permiten captar regularidades

en la formación de CN, que pueden examinarse más a fondo mediante un mecanismo de ocupación de *slots* activado por el núcleo del CN, que recibe el nombre de *microcontexto*. La traducción de CN también resultó ser un reto debido a los diferentes patrones de formación de términos y a la falta de sistematicidad en la descripción de los CN en los recursos terminográficos. En este sentido, son esenciales las técnicas de corpus y las pautas para el establecimiento de equivalentes interlingüísticos, que deben basarse en un profundo análisis contrastivo. En cuanto a la representación de los CN, debe abarcar las diferentes características de estos términos compuestos; es decir, su formación morfosintáctica y semántica, su estructura, sus variantes y equivalentes, y su posición en el sistema conceptual. Gracias a esta información, a la organización conceptual del recurso y a la posibilidad de acceder a los datos desde diferentes vías, el usuario puede obtener una descripción completa de los CN. La aplicación práctica de esta investigación es una sección de CN que se ha diseñado para el módulo fraseológico de EcoLexicon, una base de conocimiento terminológica sobre el medio ambiente.

PALABRAS CLAVE: compuesto nominal, lingüística de corpus, microcontexto, término compuesto, terminología, traducción