

Terminographic Work in Translation: exploitation of a multilingual corpus of wellness and beauty tourism (Spanish-English/French/Italian)

Cristina Castillo Rodríguez
University of Málaga, SPAIN
ccastillor(a)uma.es

ABSTRACT

The present thesis aims at confirming the hypothesis that today there is a need of terminological resources as documentation and terminology help for non-mother tongue tourist translation, especially in the subdomain of tourism called wellness and beauty. This hypothesis is divided into three sub-hypotheses: i) that specialized translation into non-mother tongue requires a thorough evaluation with the aim of assessing the quality of these translations; ii) that current resources for tourist-texts translation are not useful for solving certain terminological problems; and iii) that the subdomain of wellness and beauty tourism needs to be represented in a multilingual terminological dictionary.

To confirm or refute the aforementioned sub-hypotheses, the thesis is structured into four chapters. Chapter I deals with the description of the tourist sector in Spain, focusing on both its concept and its different tourist segments that have emerged from the beginning of the 21st century, paying special attention to the tourist segment called wellness and beauty. Furthermore, this chapter is also concerned with the description of online tourism promotion, the main tourist-emitting markets to Spain, and the current situation of tourist translation (above all, into non-mother tongue) as an independent subject in the degree of Translation and Interpreting in Spain. Chapter II is devoted to the process of compilation of the multilingual corpus, composed of three sub-corpora: 1) a comparable sub-corpus (TURISPA 1), formed by four components of source texts in Spanish, English, French, and Italian; 2) a professional parallel sub-corpus (TURISPA 2), whose component in Spanish constitutes the source texts, and of which the components in English, French, and Italian are translations of those previous texts extracted from promotional material (published entirely on the Internet) of the wellness and beauty segment; and 3) a semi-professional parallel sub-corpus (TURISPA 3), formed by a component of texts originally written in the Spanish language, and three other components of texts translated into English, French, and Italian carried out by undergraduate students of the final year of the degree of Translation and Interpreting. One of the main results of this chapter is the creation and compilation of the three sub-corpora and the classification of the most important mistakes detected in the two parallel sub-corpora.

The two last chapters are focused on the research line of terminology and terminography. Chapter III is devoted to the theoretical framework with regard to the research on terminology, as well as the terminographic work. Finally, in Chapter IV a protocolized methodology for the elaboration of a multilingual terminological dictionary is proposed with different phases by means of the use of the prototype called SpaTerm, developed and implemented by the Research Institute for Computational Linguistics (University of Wolverhampton). The fields contained in the database of this prototype were designed in the frame of this doctoral thesis, highlighting the field of terminological definition of the terminological units of the wellness and beauty sub-domain. That definition was elaborated

thanks to a previous establishment of conceptual relations observed in the different components of the multilingual comparable sub-corpus.

KEYWORDS: comparable and parallel corpus, corpus linguistics, multilingual corpus, specialized translation, terminology, wellness and beauty tourism.

Completion of Thesis

Place: University of Málaga, Spain

Year: 2010

Supervisor: Professor Gloria Corpas Pastor

Original Language: Spanish

El Trabajo Terminográfico en Traducción: explotación de un corpus multilingüe de turismo de salud y belleza (español-inglés/francés/italiano)

RESUMEN

La presente tesis doctoral parte de la hipótesis principal de que, en la actualidad, existe una gran necesidad de recursos terminológicos como ayuda documental y terminológica para la traducción turística inversa, especialmente en el subdominio del turismo denominado salud y belleza. Esta hipótesis general se subdivide, a su vez, en tres subhipótesis, a saber, i) que la traducción especializada inversa en el subdominio del turismo de salud y belleza precisa de una evaluación exhaustiva para valorar la calidad de estas traducciones; ii) que los recursos actuales para la traducción de este tipo de textos no son útiles para solventar determinados problemas terminológicos; y, por último, iii) que el subdominio del turismo de salud y belleza necesita ser representado en un diccionario terminológico multilingüe.

Para corroborar o refutar las subhipótesis anteriormente indicadas, la tesis se estructura en cuatro capítulos. El capítulo I se ocupa de la descripción del sector turístico en España, tanto su concepto como los distintos segmentos turísticos que han emergido desde principios del siglo XXI, prestando especial atención a la descripción del turismo de salud y belleza. Se aborda, además, la promoción turística electrónica, los principales mercados emisores de turistas a España y la situación actual de la traducción turística (sobre todo, inversa) como asignatura independiente en los planes de estudio de las titulaciones de Traducción e Interpretación. El capítulo II trata el proceso de compilación del corpus multilingüe compuesto por tres subcorpus: 1) un subcorpus comparable (TURISPA 1), formado por cuatro componentes de textos originales en español, inglés, francés e italiano; 2) un subcorpus paralelo profesional (TURISPA 2), cuyo componente español constituye los textos originales, mientras que los componentes en inglés, francés e italiano son traducciones de los anteriores extraídos del material promocional publicado en Internet del segmento salud y belleza; 3) un subcorpus paralelo semiprofesional (TURISPA 3), integrado por un componente original en español y otros tres componentes de textos traducidos al inglés, francés e italiano por alumnos de último curso de la licenciatura de Traducción e Interpretación. Uno de los principales resultados de este capítulo es la creación y compilación de los tres subcorpus y la clasificación de los errores más importantes detectados en los subcorpus paralelos.

Los dos últimos capítulos están más enfocados en la línea de terminología y terminografía. En el capítulo III se presenta el armazón teórico relativo a la investigación de la terminología, así como del trabajo terminográfico. Por último, el capítulo IV se centra en la metodología protocolizada para la elaboración de un diccionario terminológico multilingüe, para lo cual se proponen una serie de fases bien diferenciadas a través del prototipo SpaTerm, desarrollado por el Research Institute for Computational Linguistics (Universidad de Wolverhampton). Los campos contenidos en la base de datos de esta aplicación se diseñaron en el seno de esta tesis doctoral, destacando el campo de la definición terminológica de las unidades terminológicas del subdominio del turismo de salud y belleza, elaborada a partir del establecimiento de relaciones conceptuales observadas en los distintos componentes del subcorpus comparable multilingüe.

PALABRAS CLAVE: corpus multilingüe, lingüística de corpus, terminografía, terminología, traducción especializada, turismo de salud y belleza.