On the Use of Terminological Records in Specialised Translation

Federica Vezzani
federica.vezzani@phd.unipd.it
Department of Linguistic and Literary Studies
University of Padua

Giorgio Maria Di Nunzio
giorgiomaria.dinunzio@unipd.it
Department of Information Engineering
University of Padua

ABSTRACT

In this paper, we focus on the teaching of specialised translation and, in particular, on the preliminary phase of the translation process which is based on a broad and systematic work on the terminology of the micro-language considered. We present a new model of bilingual terminological record, as a digital tool supporting the process of translation of medical documents. Finally, we describe the results of a set of experiments which we have run since 2017 with two groups of students of the master's degree of the University of Padua.

PAROLE CHIAVE

Terminology, Specialised Translation, Terminological Record, Ternbase.

INTRODUCTION

In recent years, the need for interdisciplinarity has become increasingly important in the field of teaching and it represents a constant goal in all modern educational theories [1]. In this context, the use of digital tools for teaching Humanities has proved to be effective and efficient. For instance, the use of systems that automatically support the process of translation of documents [2] has become widespread in the teaching of Translation Studies [3, 4]. In particular, the digital tools that help to teach the methodology behind the translation of specialised texts [5, 6] facilitate students in understanding the complex translation process. Computer-assisted translation (CAT) tools, such as SDL Trados,1 Omega T,2 MateCat,3 etc., make the translation of a specialised text an interactive process between human and computer [7].

In this paper, we focus on the preliminary phase of the translation process which is based on a broad and systematic work on the terminology of the specialised language considered [8–10] such as medical, economic or legal language. Indeed, the study of terminology and the use of digital terminological resources are fundamental for students in order to 1) create their basis of knowledge in specialised disciplines, 2) master the terminology of a specific micro-language and 3) collect all the relevant linguistic data for both the source and target languages involved.

The collection and organisation of terminological data are performed through terminological records [11] providing linguistic information both for the term in the source language and for the translating candidate. These records are usually gathered into specific terminological databases which constitute a valid support for translators, terminologists and language professionals in general [12–14].

In this paper, we illustrate a new model of bilingual terminological record for the translation of medical documents which has been used in two courses of the master's degree in Modern Languages for the International Communication and Cooperation of the University of Padua: Computer Assisted Translation Tools and Specialised Translation. This digital terminological record is based on the TriMED model proposed in [15] and, in this work, we aim to test its usefulness both in the terminological analysis and in the translation process of such specialised texts.

1https://www.sdltrados.com

2http://omegat.org

3https://www.matecat.com

This paper is organised as follows: in Section 2, we contextualise the process of translation based on a terminological approach; then, we describe our terminological record as a digital tool for teaching specialised translation. In Section 3, we present the experiments with the students. Finally, in Section 4 we give our conclusions and some hints on future works.

TEACHING TERMINOLOGY FOR MEDICAL TRANSLATION

Technical-scientific translation concerns the process of transmission of technical information from a source language into a target language in specialised domains of human activity [16]. The main difficulties in learning, in particular, the medical translation process are the lack of medical knowledge of students and their unfamiliarity with medical terminology and phraseology [17, 18]. Amal Jammal states that "*[...] on ne peut traduire que ce que lon comprend bien*", that is to say it is necessary to understand and familiarize with the topics covered in the specialised text before starting the translation process [19].

Figure 1: Part of a TriMED terminological record for the technical term in the source language (EN) Ecchymosis and its equivalent in the target language (FR) Ecchymose.

Therefore, the translation process is usually divided into two main operations: the decoding and the subsequent transcoding of medical information [19, 20]. The first step consists in the decoding of technical terms conveying medical information in the source language. This means that, in order to understand specialised notions, students (as future professional translators) should create their terminological corpus, both in the source and target languages, gathering most of the relevant documents for the topic to be translated and becoming familiar with medical terminology. The second phase, involving the process of translation itself, concerns the transcoding of the lexical units previously identified into the target language by respecting the faithful transmission of information from a semantic viewpoint and the morphological and syntactic structure of the working language. In this context, terminological records are commonly used as tools gathering information about a term, its translating candidate and the related concept [11, 21].

TriMED TERMINOLOGICAL RECORD

The new model we propose is TriMED terminological record which was designed for language professionals and it has been used by students while learning the terminological analysis and the translation process itself [15]. This resource provides all the information necessary in order to decode and then transcode the meaning of a technical term and its correct translating candidate. It is structured around four axes of analysis:

1. Formal features;
2. Semantics;
3. Corpus;
4. References.

Regarding the formal and lexical framework of the term, our model of terminological record provides information such as: gender, spelling, pronunciation in the International Phonetic Alphabet (IPA) and other information about the etymology, such as derivation and composition of the term. Based on the WordNet resource,4 the record contains also all the nouns, verbs, adjectives, and adverbs deriving from the analysed term and which fall into the same semantic sphere.

The second section focuses on the semantic features of the term. First, we propose a definition extracted from reliable resources such as Merriam-Webster Medical Dictionary5 or MediLexicon6 especially for acronyms and abbreviations. In addition, we provide the semic analysis of the term [22] that is a methodology used in compositional semantics in order to decompose the meaning of technical terms (lexematic or morphological unity) into minimal units of meaning: the semes. Moreover, in order to evaluate the semantic behaviour, we collect the phraseology of the term by considering cases of collocations [23] and colligations [24].

In the corpus section, we provide all specialised contexts where technical terms have been extracted and then we proceed through the identification of the domain and the register of communication of the term (popular, standard and specialised). The term and its definition, therefore, take on meaning when they are connected to a specific domain: in our analysis, we identify the domain and subdomains of the text (such as surgery, pathology, pharmacology, etc). Finally, since all of this information has been extracted from different sources, we provide references to each source.

The Web application that allows users to manage and store TriMED terminological records was realised with the Shiny R package [25]. The interface was specifically designed to display the information of each field in the source and target language, as shown in Figure 1. Currently, the TriMED terminological database supports three working languages (English, French and Italian) and contains about 328 terminological records that are under review. A demo of the application is available.7

EXPERIMENTS

In this section, we present the studies that we have carried out on the use of the TriMED terminological record for teaching the translation process of medical texts both in the decoding and transcoding phase. In particular, we describe the educational applications concerning the use of terminological records by students of the master's degree course in Modern Languages for the International Communication and Cooperation of the University of Padua.

DECODING PHASE

The decoding phase of a specialised translation concerns the collection of a corpus of texts that can form the basis of preliminary knowledge in the process of translation of texts. In 2017/18, we performed an experiment with 90 students of the course on Computer Assisted Translation Tools. The objective of this experiments was to teach: i) how to correctly interpret pieces of medical information about a disease provided by a physician and ii) how to produce some new information that might help the physician to find relevant documents about the disease.8

Students were divided in groups of 3 people. We provided each group with a specialised medical query about a disease written by a physician and a scientific article (related to the query) retrieved from the PubMed database9 (one of the largest databases

---

4https://wordnet.princeton.edu

5https://www.merriam-webster.com/

6https://www.medilexicon.com/

7https://gmdn.shinyapps.io/TriMED/

8This study was part of a larger experiment that involved our participation to the Cross-Language Evaluation Forum (CLEF) task on Technologically Assisted Re- views [26]. For space reasons, we omit the details of the technical part.

9https://www.ncbi.nlm.nih.gov/pubmed/

of biomedical literature). Each group was asked to rewrite the original query by following a terminological approach and produce an alternative that might have helped the physician to find additional relevant documents about the disease. Finally, they were asked to translate the scientific article provided. In order to complete the task, we asked to the students to extract all the technical terms and to fill-in the TriMED terminological records. In this way, the students were able to collect all the terminological data related to the extracted terms and become familiar with the medical field before starting the translation process. The results of the experiment showed that, by means of the TriMED record, the students were able to decode correctly the specialised text and produce some new medical information that had the same quality of the original pieces of information provided by the expert [27]. For example:

- Original Query: PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer
- Alternative Query: Positron emission tomography Computed tomography for the evaluation of mediastinums lymph nodes in doubtful cancer affected people with removal NSCLC.

At present, we are running a second experiment with the students of the course of 2018/19.

TRANSCODING PHASE

The second step of the translation process concerns the transcoding of the information extracted from the text in the source language into the target language in a clear and faithful way. Currently, we are performing an experiment concerning the use of TriMED terminological records for teaching Specialised Translation in the medical domain to the students of the second year of the Master's Degree course.

Students work with Italian and French languages and they are asked to trans- late two documents (French to Italian and Italian to French) about the practice of vaccinations. In particular, they are translating two pedagogical dossiers for collective protection extracted from the Santé Publique - France website and the Local Social Health Unit of Padua website. Students are working on the technical terms extracted from the specialised texts, and they are compiling terminological records in order to choose the correct translating candidate.

After the manual extraction of technical terms, they proceed with the creation of a new empty terminological record from the TriMED application and then they manually fill all its fields with linguistic information extracted from reliable web sources. For each terminological record we save the creation date and the author and students have the possibility to modify and adjust the information provided. In order to avoid ambiguity from the semantic viewpoint, students have to fulfil the records with terminological data related only to the specific field or subfield of interest (in our case, medical field and in particular vaccination subfield).

At the moment, terminological records are proving to be a valid support for the students in the transmission of medical information. Moreover, TriMED has a user-friendly interface and this allows quick data entry and consultation.

CONCLUSIONS

In this paper, we focused on the preliminary phase of the specialised translation process which is based on the study of technical terms and their analysis by means of terminological records. In particular, we presented our methodology for the translation of medical texts through TriMED terminological record which is structured in order to provide a useful support for students (as future translators) in the two major operations involved in the translation process: the decoding and the subsequent transcoding of medical information.

We performed two experiments with the students of the master's degree course of the University of Padua and the model we proposed proved to be valid for the: i) creation of students' basis of knowledge in medical domain, ii) mastery of the terminology and iii) collection of all the relevant linguistic data for both the source and target languages. As future work, we are working on a method in order to import the terminological data obtained through the TriMED application on an assisted translation tool like SDL Trados.

REFERENCES

[1] Rossi, P.G., Biondi, S.: Interdisciplinarità. Education Sciences & Society 5(1) (2014)

[2]   Hutchins, W.J.: Machine translation over fifty years. Histoire epistémologie langage 23(1) (2001) 7–31

[3]   Toury, G.: Descriptive Translation Studies and beyond: revised edition. Volume 100. John Benjamins Publishing (2012)

[4]   Gambier, Y., Van Doorslaer, L.: Handbook of translation studies. Volume 1. John Benjamins Publishing (2010)

[5]   Scarpa, F.: La traduzione specializzata. Un approccio didattico professionale. Hoepli (2008)

[6]   Musacchio, M.T.: Terminology and specialized translation. Textus online only. 12 (1999), N. 2, 1999 12(2) (1999) 1000–1021

[7]   Hutchins, J.: Machine translation and human translation: in competition or in complementation. International Journal of Translation 13(1-2) (2001) 5–20

[8]   Cabré, M.T.: Terminology and translation. Gambier, Yves & Luc van Doorslaer (eds.) (2010) 356–365

[9]   Wright, S.E., Budin, G.: Handbook of terminology management: application- oriented terminology management. Volume 2. John Benjamins Publishing (2001)

[10]   L'Homme, M.C.: La terminologie: principes et techniques. Pum (2004)

[11]   Gouadec, D.: Terminologie: constitution des données. AFNOR gestion. AFNOR (1990)

[12]   L'Homme, M.C.: Dicoinfo, dictionnaire fondamental de l'informatique et de l'internet. Consulté à l'adresse http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search. cgi (2009)

[13]   L'Homme, M.C., Lanneville, M.: Dicoenviro. dictionnaire fondamental de l'environnement. Consulté à l'adresse http://olst.ling.umontreal.ca/cgibin/dicoenviro/search. cgi (2014)

[14]   Johnson, I., Macphail, A.: Iate–inter-agency terminology exchange: Development of a single central terminology database for the institutions and agencies of the european union. In: Proceedings of the Workshop on Terminology resources and computation, LREC 2000 Conference. Ath`enes, Gr`ece. (2000)

[15]   Vezzani, F., Di Nunzio, G.M., Henrot, G.: Trimed: A multilingual terminological database. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation LREC 2018, Miyazaky, Japan, May 7-12, 2018.

[16]   Pinchuck, I.: Scientific and technical translation. Westview Press (1977)

[17]   Wakabayashi, J.: Teaching medical translation. Meta 41(3) (1996) 356–365

[18]   M., R.: La terminologie médicale et ses probl`emes. Tribuna, Vol. IV, n. 12 (2003)

[19]   Jammal, A.: Une méthodologie de la traduction médicale. Meta: journal des traducteurs/Meta: Translators' Journal 44(2) (1999) 217–237

[20]   Rouleau, M.: La langue médicale: une langue de spécialité à emprunter le temps d'une traduction. TTR: traduction, terminologie, rédaction 8(2) (1995) 29–49

[21]   Francoeur, A.: La fiche terminologique, entre théorie et pratique. Langues et linguistique (35) (2015) 24–39

[22]   Rastier, F.: Sémantique interprétative. Formes sémiotiques. Presses universitaires de France (1987)

[23]   Firth, J.R.: A synopsis of linguistic theory, 1930-1955. Studies in linguistic analysis (1957)

[24]   Sinclair, J.: Reading concordances: An introduction. Pearson Longman (2003)

[25]   CHANG, W.: Shiny: Web Application Framework for R. (2015) R package version 0.11.

[26]   Di Nunzio, G.M., Beghini, F., Vezzani, F., Henrot, G.: An interactive two- dimensional approach to query aspects rewriting in systematic reviews. IMS unipd at CLEF ehealth task 2. In: Working Notes of CLEF 2017 - Conference and Labs of the Evaluation Forum, Dublin, Ireland, September 11-14, 2017. (2017)

[27]   Di Nunzio, G.M., Vezzani, F.: Using R markdown for replicable experiments in ev- idence based medicine. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction - 9th International Conference of the CLEF Association, CLEF 2018, Avignon, France, September 10-14, 2018, Proceedings. (2018) 28–39