

Teaching Specialised Translation Through Corpus Linguistics: Translation Quality Assessment and Methodology Evaluation and Enhancement by Experimental Approach

Natalie Kübler, Alexandra Mestivier et Mojca Pecman

Volume 63, numéro 3, décembre 2018

Traductologie de corpus : 20 ans après

URI : <https://id.erudit.org/iderudit/1060174ar>

DOI : <https://doi.org/10.7202/1060174ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Kübler, N., Mestivier, A. & Pecman, M. (2018). Teaching Specialised Translation Through Corpus Linguistics: Translation Quality Assessment and Methodology Evaluation and Enhancement by Experimental Approach. *Meta*, 63(3), 807–825. <https://doi.org/10.7202/1060174ar>

Résumé de l'article

La traduction spécialisée est, de nos jours, plus que jamais liée aux évolutions dans la communication spécialisée et aux changements affectant les discours spécialisés qui sont à la fois rapides et importants. Dans un tel contexte, la nécessité de concevoir des méthodes pour assurer la qualité de la traduction spécialisée et une formation en traduction spécialisée efficace est évidente. Le renouvellement constant des paradigmes de connaissances a plusieurs conséquences : le nombre de termes nouvellement forgés ne cesse de croître, les structures lexico-phraséologiques subissent les modifications et les conventions discursives s'adaptent aux nouveaux besoins communicatifs. Cette situation oblige les enseignants en traduction spécialisée à former les futurs traducteurs de façon qu'ils puissent acquérir les compétences nécessaires pour s'adapter rapidement à tous ces changements. Les outils offerts par la linguistique de corpus permettent d'accéder aux discours émergents ainsi qu'aux termes, collocations et structures d'une langue en mutation. Quant aux méthodes d'exploitation efficace des corpus dans les cours de traduction, elles nécessitent d'être explorées et améliorées. Dans cette étude, nous présentons une méthodologie d'enseignement de la traduction spécialisée conçue spécifiquement avec cet objectif en vue. Cette méthodologie s'appuie sur la linguistique de corpus, l'analyse terminologique, la collaboration avec des experts et l'analyse quantitative de la qualité des traductions des apprenants pour mieux cibler les difficultés spécifiques qu'ils rencontrent en traduction spécialisée et pour leur proposer les démarches appropriées pour les résoudre.

Teaching Specialised Translation Through Corpus Linguistics: Translation Quality Assessment and Methodology Evaluation and Enhancement by Experimental Approach

NATALIE KÜBLER*

Université Paris Diderot, Paris, France
nkubler@eila.univ-paris-diderot.fr

ALEXANDRA MESTIVIER*

Université Paris Diderot, Paris, France
avolansk@eila.univ-paris-diderot.fr

MOJCA PECMAN*

Université Paris Diderot, Paris, France
mpecman@eila.univ-paris-diderot.fr

RÉSUMÉ

La traduction spécialisée est, de nos jours, plus que jamais liée aux évolutions dans la communication spécialisée et aux changements affectant les discours spécialisés qui sont à la fois rapides et importants. Dans un tel contexte, la nécessité de concevoir des méthodes pour assurer la qualité de la traduction spécialisée et une formation en traduction spécialisée efficace est évidente. Le renouvellement constant des paradigmes de connaissances a plusieurs conséquences: le nombre de termes nouvellement forgés ne cesse de croître, les structures lexico-phraséologiques subissent les modifications et les conventions discursives s'adaptent aux nouveaux besoins communicatifs. Cette situation oblige les enseignants en traduction spécialisée à former les futurs traducteurs de façon qu'ils puissent acquérir les compétences nécessaires pour s'adapter rapidement à tous ces changements. Les outils offerts par la linguistique de corpus permettent d'accéder aux discours émergents ainsi qu'aux termes, collocations et structures d'une langue en mutation. Quant aux méthodes d'exploitation efficace des corpus dans les cours de traduction, elles nécessitent d'être explorées et améliorées. Dans cette étude, nous présentons une méthodologie d'enseignement de la traduction spécialisée conçue spécifiquement avec cet objectif en vue. Cette méthodologie s'appuie sur la linguistique de corpus, l'analyse terminologique, la collaboration avec des experts et l'analyse quantitative de la qualité des traductions des apprenants pour mieux cibler les difficultés spécifiques qu'ils rencontrent en traduction spécialisée et pour leur proposer les démarches appropriées pour les résoudre.

ABSTRACT

In the current context of rapid and constant evolution of global communication and specialised discourses, the need for devising methods for ensuring both high quality levels of specialised translation and successful translation training is becoming a true challenge. Steady renewal in knowledge paradigms leads to an increase in term coinage, modifications in lexical and phraseological patterns, and accommodations in discourse conventions. This situation requires teachers in specialised translation to train future translators to develop the skills meant to help them adapt rapidly to change. The tools brought by corpus linguistics offer access to the language-in-the-making and continuously emerging knowledge fields. However, methods for their efficient exploitation in translation

classes can still be improved. In the current study, we present the translation-teaching framework devised specifically for such contexts. It is based on corpus linguistics, terminology management, collaboration with experts, and the quantitative analysis of the quality of finished translations, which can then, in turn, be used to improve the overall framework and to provide research material on specialised translation problems.

RESUMEN

La traducción especializada está, hoy en día, más vinculada que nunca al desarrollo de la comunicación especializada y a los cambios que afectan a los discursos especializados, que son a la vez rápidos e importantes. En este contexto, es evidente la necesidad de diseñar métodos que garanticen la calidad de la traducción especializada y la eficacia de la formación en traducción especializada. La constante renovación de los paradigmas del conocimiento tiene varias consecuencias: el número de nuevos términos forjados sigue creciendo, las estructuras léxico-fraseológicas sufren modificaciones y las convenciones discursivas se adaptan a las nuevas necesidades comunicativas. Para los profesores de traducción especializada, esta situación les obliga a formar a los futuros traductores para que puedan adquirir las competencias necesarias para adaptarse rápidamente a todos estos cambios. Las herramientas ofrecidas por la lingüística de corpus proporcionan acceso a los discursos emergentes, así como a los términos, las colocaciones y las estructuras de un idioma cambiante. Es necesario explorar y mejorar los métodos para su uso efectivo en los cursos de traducción. En este estudio presentamos una metodología para la enseñanza de la traducción especializada diseñada específicamente con este objetivo. Esta metodología se basa en la lingüística de corpus, el análisis terminológico, la colaboración con expertos y el análisis cuantitativo de la calidad de las traducciones especializadas de los estudiantes de traducción, con el fin de identificar mejor las dificultades específicas que encuentran y proponer ejercicios adecuados para resolverlas.

MOTS-CLÉS/KEYWORDS/PALABRAS CLAVE

formation en traduction spécialisée, corpus spécialisés et annotés, typologie des erreurs de traduction, analyses quantitatives et qualitatives, construction du matériel pédagogique specialised translation teaching, specialised and annotated corpora, translation error typology, quantitative and qualitative analysis, learning material design

formación en traducción especializada, corpus especializado y comentado, tipología de errores de traducción, análisis cuantitativos y cualitativos, construcción de materiales didácticos

1. Introduction

Corpus linguistics for specialised translation teaching opens up new possibilities for analysing the different aspects of language use and production as well as for testing the adequacy of the various methods for training translation learners. In this study, we present how these possibilities were explored and exploited within a translation-teaching framework developed at Paris Diderot University.

Conceived of as a terrain for exploratory research on translation problems and issues related to developing translation skills among students, this teaching-oriented initiative is the result of fifteen years of collaborative research work in specialised translation, with emphasis on corpus linguistics, terminology management, and teamwork with experts. The association of these three related fields of knowledge and practice, and their implementation as stages prior to translation tasks, is a cutting-edge feature of this multidisciplinary teaching framework.

The core methodology consists in using corpora both for improving the translation process and for evaluating the output. Every year, our MA students in specialised translation¹ perform two translation tasks – the first without the use of corpora, and the second using comparable specialised corpora built by themselves for the purposes of the task. The translations produced are evaluated and annotated using a translation error typology (Castagnoli, Ciobanu, *et al.* 2011) developed within the framework of the MeLLANGE (Multilingual eLearning in LANGuage Engineering) European Leonardo da Vinci Project (2004-2007). They are then compiled into two translation sub-corpora (SP-TRANS1 and SP-TRANS2) according to the two conditions of production, which allows us to evaluate the impact of corpus use on specialised translation. The most salient and frequent types of translation errors, identifiable through quantitative analysis, are thereafter a focus for further improvements of our teaching methodology.

Studies on translation training are enlightening, among others (Bowker and Bennison 2003; Pearson 2003; Castagnoli, Ciobanu, *et al.* 2011; Babych, Hartley, *et al.* 2012; Kageura, Hartley, *et al.* 2015²). They all highlight the fact that translation is among the most complex cognitive processes, one calling for vast knowledge and various competences: language competences, linguistic awareness, information retrieval skills, corpus linguistics, terminology processing, etc. In this era of fast information retrieval, intercultural communication needs, and increasing demand for drafting multilingual texts, there is an urgent need to develop efficient methods and on-task solutions for training future translators.

Several models have been proposed for the systematisation of translation competences and the harmonisation of translation training, namely the model established by the PACTE research group (PACTE 2000), Susanne Göpferich's (2009) model, and the European Master's in Translation (EMT) model (EMT expert group, 2009).³ Our MA programme adopted the latter in 2009, when it joined the EMT network. We used this occasion to revisit the curriculum of our MA in specialised translation, bearing in mind the reference framework for the competences applied to translation as a profession. As stated in the 2017 edition of the EMT framework, our approach to translation

is based on the premise that “translation” is a process designed to meet an individual, societal or institutional need. It also recognises that it is a multi-faceted profession that covers the many areas of competence and skills required to convey meaning (generally, but not exclusively, in a written medium) from one natural language to another, and the many different tasks performed by those who provide a translation service.⁴ (Toudic and Krause 2017: 4)

The complex project described in the present paper addresses, consequently, a vast number of competences ranging from linguistic, sociolinguistic, textual, and intercultural to information mining and technological skills. The focus of the present study is, however, placed on the exploration of ways to improve translation training by devising corpus-based tasks specifically tailored to address difficulties encountered by students during specialised translation.

The article is structured as follows: first, we briefly review current advances in the field of corpus linguistics as applied to specialised translation teaching. Next, the translation-teaching framework and the methodology used for compiling an error-annotated translation corpus are presented. The section that follows focuses

on the quantitative analysis of the learner translation corpora. Finally, we provide examples of classroom activities we have created in an attempt to help students learn to find corpus solutions to the most common translation difficulties they encounter.

2. Specialised translation teaching as a field of research: recent advances and everlasting challenges

Together with second language teaching, specialised translation teaching is a major field of investigation in applied linguistics. The knowledge about translation that has accumulated over time is remarkable and is the result of various approaches to this practice, including theoretical (Ladmiral 1994), historical and sociological (Ladmiral 2014), and pragmatic approaches, which, in turn, can be oriented towards professional translation (Froeliger 2013), code analysis of translation-produced texts (or “third-code,” see Baker 1989) or specialised translation issues (Rogers 2015). In general, studies in the field of translation offer a rich reservoir of knowledge about translation processes, translation teaching techniques, and aspects of translated texts (Baker 1999; Olohan and Baker 2000; Puurtinen 2003; Olohan 2004; Mauranen 2007; Frankenberg-Garcia 2009).

However, works reporting experimental results of corpus use in translation teaching tend to be more recent. Bowker and Bennison (2003), for instance, discuss the results they obtained by using a Student Translation Tracking System, a tool allowing students to evaluate their own productions. Pearson (2003) on the other hand, explores the usefulness of a small parallel corpus in a classroom for specialised translation tasks, which gives students the opportunity to compare their productions with those of professionals. The objective of Pearson’s teaching method is to draw students’ attention to translational strategies employed by professionals. Loock, Mariaule, *et al.* (2013) show how the outcomes of inquiries into the use of corpora for translation can be usefully exploited for pedagogical purposes. Within the framework of the MeLLANGE project, Castagnoli, Ciobanu, *et al.* (2011) propose a method inviting students to query multilingual annotated corpora through several series of exercises in order to train them to identify various strategies for resolving specific translation issues.

Recently, the MNH-TT (Minna no Hon’yaku for Translation Training) project developed at the University of Tokyo has provided further evidence of a need for increased sophistication in translator training schemes (Babych, Hartley, *et al.* 2012; Kageura, Hartley, *et al.* 2015). The specific aim of this project is to provide a platform whereby trainee translators can gain competences not only in “translation” in the narrow sense but also in the undertaking and management of translation projects. As translation activities are increasingly becoming project-based and collaborative while making regular use of online environments, this project allows students to gain useful reflexes from early exposure to such a working model.

At Paris Diderot University, a continuously improved teaching framework has been the subject of experimentation for fifteen years (Kübler 2003), during which the teaching of specialised translation has resorted to corpus linguistics and a multidisciplinary approach to the translation process. Since 2013, this framework has integrated an additional component: the systematic collection of translations produced

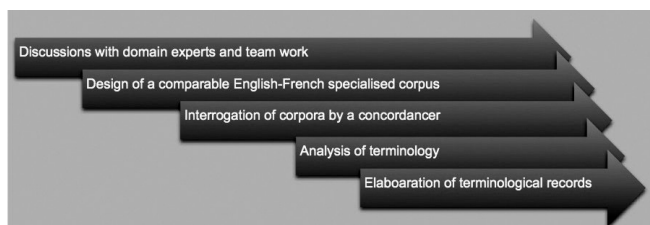
with and without corpora, followed by quantitative analysis of translation errors (Kübler, Mestivier, *et al.* 2015⁵; Kübler, Mestivier, *et al.* 2016).

3. An overall description of the teaching framework

The translation teaching framework project we present in this study is the outcome of long-term research into translation issues, translation teaching techniques, and innovation in this area of applied linguistics. The framework is a part of the curriculum for first year MA studies in specialised translation, that is, students enrolled in a Master of Applied Languages in the EILA (Études Interculturelles de Langues Appliquées) department of Paris Diderot University. Its main feature consists in exploiting the advantages of not only related areas of knowledge but also a multidisciplinary approach to specialised texts that allows students to immerse themselves in specialised discourse before embarking upon translation, namely through the practice of corpus linguistics, terminology management and collaborative teamwork with domain experts. Figure 1 summarises the parallel stages that translation learners go through during the first term, before moving on to translation tasks in the second term.

FIGURE 1

Stages preliminary to the translation task



Since the project is conducted in collaboration with the Earth and Planetary Sciences (EPS) department of Paris Diderot University – UFR Sciences de la Terre et de l’Environnement et des Planètes (STEP) – translation tasks are based on scientific articles covering various aspects of this field. The selected texts are recently published English texts culled from high-impact, peer-reviewed journals, such as *Nature*, *Science*, and *Earth and Planetary Science Letters*. The students from both the EILA and STEP departments attend joint classes during the first semester (generally 6 classes out of 12, amounting to 12-hour teamwork). STEP students need to improve their linguistic competences in English and, at the same time, discover the conventions associated with a given genre, namely scientific articles. They have a separate task: preparing, in English, oral presentations of all the articles that EILA students are tasked with translating.⁶

At each step, the students are guided in their discovery of the related fields of study (cf. Figure 1) in order to develop a set of related competences that will equip them to achieve the translation quality required in a professional setting. Specifically, they discover how to:

- compile a comparable corpus
- use concordances to query and effectively exploit corpora
- conduct terminological analysis
- create terminological records
- collaborate with experts.

This facilitates their comprehension of a specialised domain and helps them understand several aspects of working with specialised discourse, namely terminology, phraseology, syntax, and register.

Terminological and phraseological analysis is used as a base for identifying the domain's key concepts and the linguistic means to express them. It leads to the construction of dictionary entries (terminological records), which are compiled in the ARTES⁷ online database whose purpose is the provision of essential language resources for all students during the translation process (Pecman and Kübler 2011). The ARTES database offers scaffolding for students in the collection and arrangement of knowledge-rich information. Students are given the task of collecting and producing comprehensive definitions for retrieved terms, providing useful contexts as examples, identifying and recording collocations, synonyms, cognate terms, and adding all the necessary remarks or observations for the efficient use of terms. Each student constructs five bilingual terminological records, after the terminological analysis and prior to the translation task.

Students simultaneously compile a collective comparable specialised corpus by gathering research articles, PhD dissertations, online course materials, etc. The corpus is indexed with IMS Corpus Workbench in order to obtain efficient queries of textual resources with the Corpus Query Processor (CQP) language.⁸ Table 1 shows the size of the Earth and Planetary Sciences (EPS) corpora collected over recent years.

TABLE 1

The size of comparable specialised corpora collectively compiled by translation learners

	English corpus	French corpus
2013-2014	12.4 million words	27.8 million words
2014-2015	12.9 million words	50.0 million words
2015-2016	9.3 million words	36.6 million words
2016-2017	7.9 million words	28.0 million words
2017-2018	11.8 million words	55.0 million words

The French corpora are generally larger than the English corpora as they contain fewer scientific articles and more PhD theses. Scientific articles are overwhelmingly published in English even by French researchers. This lack of resources is compensated by the collection of PhD theses written in French.

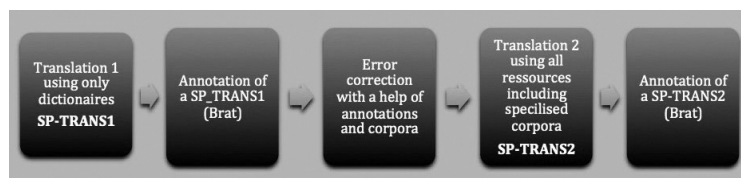
From a theoretical perspective concerning learning activities, the training methodology we devised is conceived on the basis of (inter)active pedagogy (Bonwell and Eison 1991) as it consists in incorporating active learning in small groups for addressing, analysing, and understanding the texts on specific subjects before embarking on a translation. This is all the more so as our classes tend to place emphasis on students' explorations of their difficulties and on how corpus use can help to solve them.

4. Designing an annotated translation learner corpus

Every year, two translation learner sub-corpora are compiled by gathering two sets of translations produced by our students (Figure 2).

FIGURE 2

Stages in compiling an annotated translation learner corpus



To produce translations in SP-TRANS1, students are allowed to use a limited set of dictionaries and term bases:

- TERMIUM Plus⁹
- IATE: the InterActive Terminology for Europe¹⁰
- Le grand dictionnaire terminologique¹¹
- Merriam-Webster¹²
- Word Reference¹³
- The Reverso Dictionary¹⁴
- The Pons English-French dictionary.¹⁵

Between SP-TRANS1 and SP-TRANS2, there is usually a four-week hiatus during which we annotate translations gathered in SP-TRANS1 and provide the following types of teaching input:

- a) Several teaching activities revolving around corpus querying in order to identify equivalent terms in the target language and ways to couch terms in phraseology that is acceptable for scientific discourse in the target language;
- b) Teaching input revolving around a presentation of the most frequent translation errors. As shown in Sections 6.1-6.5, we have devised various teaching activities to raise students' awareness on how corpus use can help them to avoid certain types of recurrent translation errors;
- c) Finally, access given to students to the annotated version of their initial translation (during the four-week period between SP-STRANS1 and SP-TRANS2). Using our feedback accompanying the annotation, in addition to our full support during classes, students produce a second version of the same translation, in which each segment of modified text is followed by an explanation of how they arrived at a better translation solution (including particular corpus queries, concordance screenshots, etc.).

For SP-TRANS2, students may use any resource available (including those they used for SP-TRANS1) and are strongly encouraged to make use of the EPS corpora they collectively compiled. For SP-TRANS2, specialised dictionaries and term databases remain a very precious primary resource to which the students resort to find equivalent terms. The corpus is used as a supplemental, more advanced resource when students are unable to find an equivalent term in term bases or when they wish to ascertain that the terms are couched in phraseology that is acceptable for target language scientific discourse. For each translation segment, students must provide a

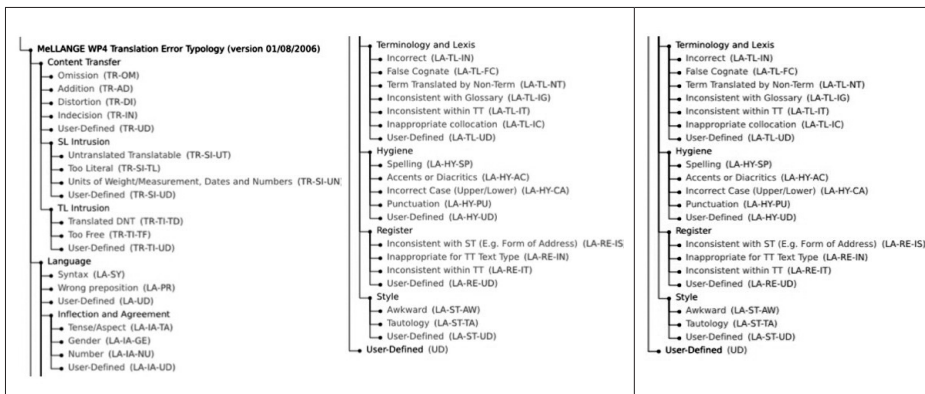
comment on the specific corpus queries or Internet searches that led them to the best translation solutions.

SP-TRANS1 and SP-TRANS2 translations are uploaded on a *brat*¹⁶ server and annotated using the MeLLANGE error typology¹⁷ (Castagnoli, Ciobanu, *et al.* 2011). *Brat* is an online environment for collaborative annotation chosen for the project for its many advantages. As a user-friendly online environment, it allowed us to view and compare annotations, give students easy access to their annotated translations, and provide them with easily understood feedback, using the comment function (Figure 4). *Brat* also offers a concordancing function that allows users to find examples of a particular type of translation error and to query annotators' comments or the translation text itself. *Brat* does not currently offer an in-built statistical analysis tool, but we decided that this was not a major drawback since the stand-off format under which *brat* annotations are stored can be easily exported and straightforwardly analysed (Section 4).

The MeLLANGE error typology was created in order to describe specific translation error phenomena rather than make any judgment on the quality of translation. It does not provide categories for encoding the degree of a translation error. The typology is based on the fundamental distinction between “content-transfer errors” and “language errors” (Figure 3). These two categories are further divided into sub-categories such as “source language intrusion” or “terminology and lexis,” which, in turn, include more specific errors such as “too literal” or “incorrect collocation.” The hierarchy also includes, at all levels, “user-defined” annotations. These “user-defined” annotations can be used either to signal a type of error not provided in the annotation scheme (which could eventually lead to a revision of the annotation scheme) or to mark highly successful translation solutions.

FIGURE 3

The MeLLANGE error typology

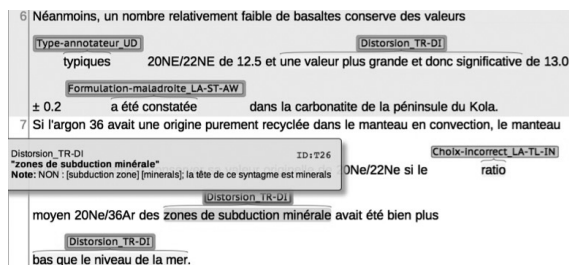


Using *brat*, each error is marked with an error code that is attached to words, phrases or entire sentences (several errors can be attached to the same lexical or textual unit), and feedback to the students is provided using the comment function (Figure 4).

The annotation of errors is conducted by five teachers/researchers (including the authors of the article). We begin by agreeing on a common procedure for error annotation: we start by all annotating the same translation sample, then compare and

discuss our annotations. However, given the fact that the annotation process is time-consuming (involving, as it does, consultation of specialised dictionaries and terminological databases, corpus querying to check terminology and phraseology, etc.) we did not take, at this stage of the project, inter-annotator agreement further, in order to be able to assess it statistically. This limitation of our study is intended to be the focus of the next developments for the increased efficiency and overall results adequacy of our methodology.

FIGURE 4
Error annotation using *brat*



The next section presents the results of the statistical analysis we conducted to determine which linguistic units and language features our students find particularly problematic or difficult when conducting translation tasks.

5. Contrastive analysis of annotated translation corpora (SP-TRANS1 and SP-TRANS2)

In Kübler, Mestivier, *et al.* (2015) and Kübler, Mestivier, *et al.* (2016), we attempted to quantify the improvements that students achieve by using the EPS corpora. In this paper, we would like to focus on illustrating how corpora can be efficiently used to avoid translation errors specifically found in learners' specialised translation productions (Section 6). However, before moving on to the teaching input we made available to students to improve their translations, we will briefly present the statistical data discussed in our previous studies.

In Kübler, Mestivier, *et al.* (2015) we compared normalised (per million words) frequencies of each error type in SP-TRANS1 and SP-TRANS2 to find out whether the error frequency decreases with the use of corpora during translation. For this study, we used translations produced over two academic years (2013-2015), that is, 108 student translations of 32 research articles compiled into a corpus containing 63 058 tokens (27 888 for SP-TRANS1 and 35 170 for SP-TRANS2).

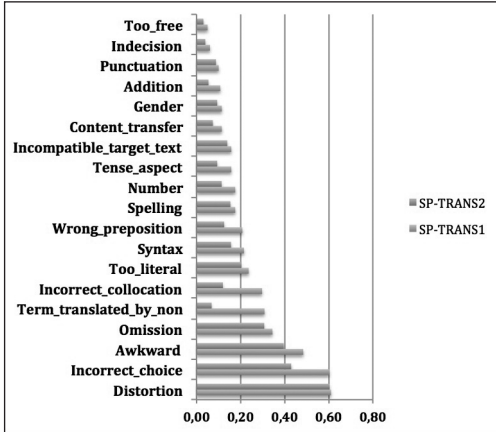
In Kübler, Mestivier, *et al.* (2016), we used a sub-corpus of student translations carried out during the 2014-2015 academic year. It comprised 55 student translations of 14 research articles. This 37 324 token corpus was also divided between SP-TRANS1 (15 311 tokens) and SP-TRANS2 (22 013 tokens). In this contribution, we adopted a textometric approach (cf. Fleury and Zimina 2014) and measured the specificities of each type of error in the two sub-corpora.

In both studies we observed that:

- a) most errors, with few exceptions, are significantly lower in the translations conducted with the help of comparable specialised corpora (Figure 5).

FIGURE 5

Normalised frequencies for error types in SP-TRAN1 and SP-TRAN2



In other words, most translation difficulties were better handled by students when translating with the help of comparable corpora. During the 2016-2017 academic year, we also required students to explicitly state for each individual sentence in their translation whether they searched the corpus or not, and we found that they used the corpus for 68% of their translation units.

- b) The most frequent errors that students manage to avoid using corpora are, respectively: term translated by a non-term, incorrect collocation, incorrect choice, wrong preposition, etc.

This confirmed our initial hypothesis that our training should focus on these particular areas and include activities on identifying terms and studying their semantic prosody (examples provided in Section 6.1), translating collocations (examples provided in Section 6.2) and mastering collocations in second-language writing (examples provided in Section 6.3).

The types of errors for which no significant or merely limited improvement was found were: distortion, awkward formulation, too literal, incompatible with target text type, etc. Even though the method we used proved to be not particularly efficient in these cases, annotating the corpora with the MeLLANGE error scheme allowed us to identify major ‘areas of difficulty’ encountered by our students – an interesting result in itself – and to attempt to find teaching solutions. For instance, many cases of “distortion” errors are due to an erroneous analysis of complex noun phrases (NPs) and, more particularly, an erroneous identification of the head in complex NPs. We therefore decided to include an activity aimed at making students aware of this incorrect analysis and the impact it has on the quality of translations (examples provided in Section 6.4). Similarly, by discussing with students the examples of the “register errors” type, we aim to make the students aware of scientific register and genre conventions (examples provided in Section 6.5).

6. Analysing and exploiting corpus data in the classroom

In the following sub-sections, we provide several examples of classroom activities addressing the most salient error types listed above. We explain each set of examples briefly, omitting the full statement that introduces each activity in the classroom.

6.1. Using the corpus to identify a term, distinguish it from its general language homonym, and observe its specificities: collocational pattern, semantic preference, etc.

A prime example is the terminological analysis of the term *constraint* and the associated verbal term, *to constrain*, in geological modelling. In most cases, students do not realise that they are dealing with terms and assume they have the same meaning and function as the corresponding lexical items used in English for General Purposes (EGP), that is, something that limits or restricts someone or something. A few examples of translated texts provided in Table 2 prove that students fail to identify the terms *constraint/constrain* and use the general language equivalent or an equivalent paraphrase having a negative semantic prosody.

TABLE 2

Examples of erroneous translations involving the terms *constraint* or *to constrain*

Samples taken from source texts	Students' translations of samples
This Archean age for recycled oceanic crust also provides key constraints on the length of time that subducted material can survive in the mantle [...]. ¹⁸	Cependant l'hypothèse de l'origine archéenne de cette croûte océanique recyclée présente des problèmes majeurs en ce qui concerne le temps pendant lequel [...].
If the evolution of Hadean melts was similarly constrained , then our results imply that [...]. ¹⁹	Si l'évolution des magmas Hadéens était contrainte aux mêmes limitations , alors nos résultats impliquent [...].
[...] but the pathways for noble gas subduction have been poorly constrained [...]. ²⁰	[...] mais les trajectoires permettant la subduction des gaz rares ont été peu freinées/restreintes [...].

A quick study of the specific collocations of this term in the EPS English sub-corpus (see Table 3 below) could, nevertheless, draw one's attention to the fact that in this particular field, *constrain/constraint* have a distinct meaning – namely, (a condition) restricting the space of possible values of a parameter – and a positive connotation.

TABLE 3

Most frequent collocations of the terms *to constrain* and *constraint* in the EPS corpus

VB collocates for <i>constraint</i>	ADJ collocates for <i>constraint</i>	ADV collocates for <i>to constrain</i>
provide constraints place constraints obtain constraints give constraints apply constraints	important constraints additional constraints observational constraints key constraints strong constraints useful constraints	poorly constrained well constrained directly constrained tightly constrained strongly constrained precisely constrain independently constrained adequately constrain

Along with other support verbs, such as *to impose* and *to place*, *constraint* is very frequently associated with the verb *to provide*, which has a very marked positive semantic prosody and which is very rarely associated with *constraint* in English for General Purposes (EGP). Similarly, the most prominent adjectival collocates associated with *constraint* are evaluative adjectives such as *new*, *strong*, *important*, *additional*, and *key*. The steps through which we invite students to discover these terms are:

- a) searching for the most frequent verbal collocations of the term *to constrain* in the EPS English sub-corpus
- b) confirming the positive semantic prosody of the verb *to provide*, both in the EPS English sub-corpus and in EGP reference corpora (for example, BNC and the CoCA)
- c) searching for the most frequent adjectival collocations of the term *constraint* and the most frequent adverbial collocations of the verb *to constrain*
- d) gaining the awareness that some of these collocations would be unthinkable in EGP (for example, *isotopic constraint*, *strong constraint*, *poor constraint*, *petrological constraint*, *tight constraint*, etc.) by querying an EGP reference corpus
- e) providing a tentative definition of the term in the field of geological modelling
- f) finding and validating the equivalent term in the EPS French corpus: contrainte and contraindre
- g) studying the collocational behaviour of the equivalent French terms
- h) proposing a better translation for the sentences given in Table 2.

We believe that taking the students through a full process of term identification, the associated terminological and phraseological analysis, and finally translation, along with a step-by-step demonstration of how to use the concordancer and how the corpus can be used to answer each question, whenever possible, will furnish them with the necessary tools for their future careers as translators.

6.2. Using the corpus to translate a collocation

Since “incorrect collocation” is one of the most frequent translation errors that students make and, at the same time, one of the error types which could be the most easily avoided by using the corpus, several classroom activities revolve around identifying better collocation candidates using the corpus.

For instance, given an example such as (1), for which the inappropriate translation (1a) was proposed, students are invited to search the corpus for the co-occurrence of the two patterns *salin* ... fluid(s)* in the French EPS corpus, using, for instance, the following CQP query: “fluides?” []{1,4} “salinité.” A rapid analysis of the concordance fragment given in Figure 6 leads them to propose an improved translation such as the one given in (1b).

- 1) *The Cl enrichment is widely attributed to the addition of slab-derived Cl in **high salinity fluids**.*

(see note 20)

- a) L'enrichissement en chlore est majoritairement attribué à l'ajout de chlore venant des plaques dans des **fluides riches en salinité**.
- b) L'enrichissement en chlore est majoritairement attribué à l'ajout de chlore provenant des plaques dans des **fluides de salinité élevée**.

FIGURE 6

A sample of collocations **fluid*... salin*** in the French EPS sub-corpus

température sont possibles à partir de 80°C en présence de **fluides** très fortement salins et un ratio fluide/ roche élevé . Cette cri
 35d) . Ceci semble indiquer la présence dans ce secteur des **fluides** de salinités variées , ce qui impliquerait plusieurs génération
 fig.42 103 Fig .44 - Natures des **fluides** : Salinités en % p.éq . NaCl et moles % CO2 ...
 uides nous a permis de consta ter que la physico-chimie des **fluides** (salinité , proportion des cons tituants majeurs H 2 O ,
 la remontée des fluides , génère deux types de fluides , un **fluide** de forte salinité et un fluide en phase vapeur condensée de fa
 fférentiellement dans la phase vapeur qui peut condenser en **fluide** de faible salinité avant sa remontée le long de la zone de
 MREE . En comparaison , le site **Menez Gwen** , avec des **fluides** à faible salinité , possède une anomalie en europium peu acc
 querait leur appauvrissement dans les cheminées liées à des **fluides** de faible salinité . De manière générale , il est admis que
 la remontée des fluides , génère deux types de fluides . Un **fluide** de forte salinité et une phase vapeur condensée de faible sali
 Na , Ca , Mg , Fe , Ba) et des **fluides** salins (riches en K , Cl et H2O) . Les
 et 500 bars à la base du complexe filonien . Ces inclusions **fluides** présentent des salinités variables , pouvant être inférieures ou
 sus de veines de stockwerk . Cette épidote serait issue d' un **fluide** ayant une salinité proche de celle de l' eau de mer avec

Several classroom activities revolve around finding the right preposition associated with a verb, finding the right adjectival collocation for noun terms, or finding the right support verb. We illustrate this particular type of collocational problem in the following example.

6.3. Using the corpus to find a support verb

The selection of the right support verb in translation can be an obstacle for a novice translator working with language for special purposes (LSP). In general, support verbs are often difficult to master for second language learners, all the more so when working with LSPs. For instance, at the beginning of the first term, students are asked to produce an English summary of the article they work on throughout the year – the article they use to undertake a terminological analysis during the first term and from which they translate a 500-word fragment during the second term. While reading their summaries, we noticed that students tend to use a very limited number of support verbs, such as *make*, *do*, *happen*, and *lead*. In one of our classroom activities, we invite students to search the English EPS corpus to find a better choice for a support verb. Provided with examples such as the ones in column 1 of Table 4, students are asked to formulate CQP queries, allowing them to propose the improved sentences shown in column 3.

TABLE 4

Using the corpus as a second-language writing-aid tool: finding the appropriate support verb

First proposal	CQP query	Ameliorated version
When S. Anzellini and his team made their experiment with the diamond anvil cell, they observed that a fast recrystallisation happened at a lower temperature than the melting temperature.	<ul style="list-style-type: none"> ▪ [pos="V.*"][]{}{0,3} "experiment.*" ▪ [pos="V.*"][]{}{0,3} "recrystallisation.*" 	When S. Anzellini and his team performed their experiment with the diamond anvil cell, they observed that a fast recrystallization occurred at a lower temperature than the melting temperature.
A team of American scientists led a research in the Cook Islands in Polynesia [...].	<ul style="list-style-type: none"> ▪ [pos="V.*"][]{}{0,3} "research" 	A team of American scientists conducted research in the Cook Islands in Polynesia [...].

Several measures have been made on the basaltic lavas and olivine to measure the content of sulphides.	<ul style="list-style-type: none"> ▪ [pos="V.*"][]{}{0,3} "measurement.*" ▪ What is the difference between <i>measure</i> and <i>measurement</i>? 	Several measurements were performed on basaltic lavas and olivine to determine their sulphide content.
--	---	--

This type of monolingual activity helps raise students' awareness to the fact that corpora are not only a valuable translation support tool but also a potential second-language writing aid tool.

6.4. Using corpora to better analyse, understand, and translate complex noun phrases

From the very beginning of our project, we were aware that the MeLLANGE annotation scheme could be revised and improved by considering the cases where the existing error categories seemed to lack – for our pedagogical ends – a measure of granularity in the description of errors. A compelling example of this shortcoming is the case of noun phrases. Complex, long noun phrases are a means of compacting information in scientific articles and are characteristic of this genre (Pecman 2014; Mestivier-Volanschi 2015; Kübler, Mestivier, *et al.* 2015; Kübler, Mestivier, *et al.* 2016). While marking “distortion” errors, we noticed that the greater part of these errors originates in incorrect analysis and, therefore, understanding of complex noun phrases (NPs).

Consequently, we devised a classroom activity intended to increase students' awareness of this issue and to suggest ways of using comparable corpora to:

- a) confirm NP interpretation, that is, identify the head and provide a syntactic representation of the complex NP, and
- b) suggest the most appropriate ways of translating NPs.

Table 5 provides a few examples of source language sentence fragments, erroneous or inaccurate student translations, and schematic syntactic representations of complex NPs, as well as suggested corpus queries that could confirm NP analysis.

Since 2016, a more fine-grained error category, “complex NP,” was introduced in the MeLLANGE error typology to allow us to single out contexts where students pick out the wrong NP head and then to include them in future classroom activities.

TABLE 5

Examples of incorrect NP analyses leading to ‘Distortion’ type translation errors

Source text	Student translation	Error analysis and solution
[...] the vein minerals representative of first breakdown fluids [...]. (see Note 20)	*[...] les veines minérales représentatives de la première rupture des fluids [...].	<ul style="list-style-type: none"> • translate <u>première rupture des fluides</u> into English (=> <i>first breakdown of fluids</i> or <i>fluids first breakdown</i>) and compare it with the English term in the source text • look in the corpus for knowledge-rich contexts to find out which is a compound term – <i>breakdown fluid</i> or <i>first breakdown</i>

[...] the presence of atmospheric noble gases in subduction-zone serpentinites [...]. (see Note 20)	*[...] la présence des gaz nobles atmosphériques dans les zones de subduction constituées de serpentinite [...].	<ul style="list-style-type: none"> • translate <u>zones de subduction constituées de serpentinite</u> into English and explain the analysis error • consider [[subduction-zone] serpentinite] • suggestion: analyse the implicit semantic relation between <i>subduction-zone</i> and <i>serpentinites</i> by switching from NN to N prep N and finding the appropriate preposition (=> <i>serpentinites in subduction zone</i>)
[...] estimation of Fe melting temperature at ICB pressure based on static compression data [...]. ²¹	*[...] l'estimation de la température de fonte de Fe à l'ICB basée sur la compression statique des données [...].	<ul style="list-style-type: none"> • consider [[static compression] data] • suggestion: confirm that <i>static compression</i> is a term

6.5. Becoming familiar with the genre

The MeLLANGE annotation scheme allowed us to annotate register-related errors, such as “inconsistent with source text” or “inappropriate for target text type.” We devised a classroom activity aimed at raising students’ awareness of the features of the genre they are translating and, more generally, to show how corpora can answer register-related questions. This activity covers examples such as (2) below. By searching the French EPS corpus for the form vous proposed in a learner translation (2a), students may reach the conclusion that addressing the reader is not a characteristic of the scientific genre.

- 2) *Here, we report the laser-heated DAC determination of the Fe melting curve from 50 to 200 GPa [...].*

(see note 21)

- a) **Vous** trouverez ici le compte-rendu de la détermination de la courbe de fusion de 50 à 200 GPa [...].

Similarly, starting with the difficulties some of our students encounter when translating the present perfect tense, as in (3) – an example taken from an article introduction, and using the present tense, as in (3a) – we introduce the Create-A-Research-Space (CARS) model (Swales 1990) to translation students (cf. EILA students) as well as students in geology (cf. STEP students). By making students more aware of the succession of moves and steps in the introduction of scientific articles, it becomes obvious that the example in (3) serves not only to indicate a gap in previous research but also to justify the research reported in the current article. Knowledge of the rhetorical structure of scientific article introductions brings students to the conclusion that the present tense is not an appropriate translation for the present perfect tense, in this example.

- 3) *Introduction: A limiting factor in unravelling the deep noble gas cycle **has been in distinguishing** real variation in mantle noble gas signatures from ‘atmospheric contamination’ originating when basaltic magmas are exposed to sea water and the atmosphere.*

(see note 20)

- a) Un des facteurs limitant pour expliquer le cycle des gaz nobles présents en profondeur dans le manteau **consiste en une distinction** entre les variations des signatures de gaz nobles dans le manteau et la ‘contamination atmosphérique,’ qui vient de magmas basaltiques exposés à l’eau de mer et à l’atmosphère.

7. Enquiry on student perception of the effectiveness of corpus use

In order to analyse learners' appreciation of the effectiveness of corpus use for specialised translation, we conducted an enquiry comprised of a series of questions on which students can freely express their point of view. Analysing their answers helps us fully comprehend the advantages and the limitations of corpora as a tool for novice translators.

It is interesting to note the convergence of the observations made by the students (notwithstanding the significant number of learners involved) and their consistency with the results and the observations of our quantitative analysis of annotated corpora of learners' translations. The following sample of answers to our query corroborates the effectiveness of corpora for enabling better quality translations while underlining their main limitations:

- the corpora are very useful, they provide data for making appropriate translation choices,
- corpora are comparable to domain experts who can help us understand and translate a specific concept,
- designing and compiling a corpus is time-consuming,
- the systematic use of corpora can evince the part that intuition plays in translation,
- the erroneous interpretation of corpus data can lead to translation errors.

8. Conclusion

This study provides evidence of the usefulness of specialised corpora for specialised translation. Our quantitative analysis of translation corpora annotated by error types reveals a significant decrease in errors in a sub-corpus of learners' translations carried out using the corpora, when compared with the corpus of translations carried out without the help of corpora. Despite the existence of a few error categories that appear not to benefit from corpus use and corpora-related drawbacks observed during enquiry analysis – namely, students' overconfidence in corpora use and the decrease of intuitive processing – the overwhelming effectiveness of corpora for specialised translation remains undeniable.

The study also shows that the combination of corpus linguistics tools, terminology management, and collaboration with domain experts is an efficient way of training translation learners to become confident and competent translators. The multidisciplinary approach to translation is, at the same time, a compelling and robust teaching method. Annotating translation learners' productions using the error typology and conducting translation tasks under different conditions together represent a useful device for evaluating the impact of the corpus on translation and the corpus-based translation teaching methodology, and for spotting and addressing the most salient errors.

Consequently, one of the crucial features of the framework described is the incremental updating of specialised translation teaching material and methodology. The key errors and the most frequent difficulties identified through quantitative analysis in this study point to the need to place additional focus on the difficulties of comprehending, analysing, and transferring the meaning of terms, complex noun groups, and collocations. The teaching framework is thus being increasingly con-

strained and tailored to resolve specific translation problems identified in corpora of learner translations, which are then exploited for designing translation exercises.

In light of the quantitative analysis of the translation results described above, we are currently looking to complement our methodology by using specifically designed tasks to handle a variety of complex nominal groups, typically found in English LSPs, vis-à-vis French ways of expressing the same units of meaning. Our teaching framework will thus continue to evolve by integrating, systematically, into our teaching material, exercises on the most frequent and salient errors encountered in learner translations observed through learner corpora analysis.

NOTES

- * Membres du Centre de Linguistique Inter-langues, de Lexicologie, de Linguistique Anglaise et de Corpus-Atelier de Recherche sur la Parole (CLILLAC-ARP), Université Paris Diderot.
1. UFR D'ÉTUDES INTERCULTURELLES DE LANGUES APPLIQUÉES (Last update: 13 Juin 2018): *Master Langues étrangères appliquées (LEA)*. Paris: Université Paris Diderot. Visited 11 September 2018, <<http://www.eila.univ-paris-diderot.fr>>.
 2. KAGEURA, Kyo, HARTLEY, Anthony, THOMAS, Martin, *et al.* (2015): Scaffolding and guiding trainee translators: the MNH-TT platform, unpublished. *Conférence TAO-CAT-2015*. Angers, 18-20 June 2015.
 3. GAMBIE, Yves (on behalf of the EMT expert group) (2009): *Competences for Professional Translators, Experts in Multilingual and Multimedia Communication*. European Master's in Translation. Bruxelles: European Commission. Visited 18 June 2018, <http://web.archive.org/web/20160304060445/http://ec.europa.eu/dgs/translation/programmes/emt/key_documents/emt_competences_translators_en.pdf>.
 4. TOUDIC, Daniel and KRAUSE, Alexandra (on behalf of the EMT Board) (2017): *Competence Framework 2017*. European Master's in Translation. Bruxelles: European Commission. Visited 18 June 2018, <https://ec.europa.eu/info/sites/info/files/emt_competence_fw_2017_en_web.pdf>.
 5. KÜBLER, Natalie, MESTIVIER, Alexandra, and PECMAN, Mojca (2015): Étude sur l'utilisation des corpus dans l'enseignement de la terminologie et de la traduction spécialisée, unpublished. *TRELA 2015: Terrains de recherche en linguistique appliquée*. Paris, 8-10 July 2015.
 6. This collaboration also includes a design of posters illustrating the advantages of a multidisciplinary teaching approach and teamwork. The posters give rise to an annual exhibition (at the University's main exhibition room) organised every January of the university year.
 7. *Dictionnaire ARTES* (Aide à la Redaction de TExtes Scientifiques) is a dictionary-assisted writing tool for scientific communication. It is a multidomain and multilingual, terminological and phraseological database developed by the Centre de Linguistique Inter-langues, de Lexicologie, de Linguistique Anglaise et de Corpus-Atelier de Recherche sur la Parole (CLILLAC-ARP) research group as an experimental tool with a two-fold aim: to offer exploitable resources to translators and scientists obliged to write in their second language, and to provide a guideline for teaching terminology management and specialised translation in the Department of Applied Languages (Études Interculturelles de Langues Appliquées, EILA) of Paris Diderot University. The database is continuously maintained and provides worthwhile results for the researchers in the group working on electronic lexicography and lexical resource design. See CLILLAC-ARP (2015-): *Dictionnaire ARTES*. Paris: UFR d'Études Interculturelles de Langues Appliquées (Université Paris Diderot). Visited 31 August 2018, <<https://artes.app.univ-paris-diderot.fr>>.
 8. EVERT, Stefan (10 July 2005): *The CQP Query Language Tutorial*. Stuttgart: Institute for Natural Language Processing (University of Stuttgart). Visited 23 October 2017, <www.ims.uni-stuttgart.de/forschung/projekte/CorpusWorkbench/CQPTutorial/cqp-tutorial.2up.pdf>.
 9. *TERMIUM Plus** (Last update: 3 August 2018). Ottawa: Government of Canada. Visited 30 August 2018, <<http://www.btb.termiumpius.gc.ca/>>.
 10. TRANSLATION CENTRE FOR THE BODIES OF THE EUROPEAN UNION (7 December 2018): *Interactive Terminology for Europe (IATE)*. Bruxelles: European Union. Visited 8 December 2018, <<http://iate.europa.eu>>.
 11. OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE (2012-): *Le grand dictionnaire terminologique*. Québec: Gouvernement du Québec. Visited 3 September 2018, <<http://www.gdt.oqlf.gouv.qc.ca>>.

12. *Merriam-Webster.com* (2015-): Springfield: Merriam-Webster. Visited 20 May 2018, <<http://www.merriam-webster.com>>.
13. KELLOGG, Michael (1999-): *WordReference.com*. Weston. Visited 27 November 2018, <<http://www.wordreference.com>>.
14. *Dictionnaire Reverso* (2006-): Paris: Reverso-Softissimo. Visited 14 May 2019, <<http://dictionnaire.reverso.net>>.
15. *Dictionnaire en ligne PONS* (2001-): Stuttgart: PONS GmbH. Visited, 13 May 2019, <<http://fr.pons.com/traduction>>.
16. STENETORP, Pontus, PYYSALO, Sampo, TOPIĆ, Goran, *et al.* (8 November 2012): *brat: rapid annotation tool*. Version 1.3. Visited, <<http://brat.nlplab.org>>.
17. See the following for the full error typology. KÜBLER, Nathalie *et al.* (1 August 2006): *MeLLANGE WP4 Translation Error Typology*. Multilingual eLearning in Language Engineering. Paris: UFR d'Études Interculturelles et de Langues Appliquées (Université Paris Diderot). Visited 13 January 2019, <http://corpus.leeds.ac.uk/mellange/images/mellange_error_typology_en.jpg>.
18. CABRAL, Rita A., JACKSON, Matthew G., ROSE-KOGA, Estelle F., *et al.* (2013): Mass independently fractionated sulfur isotopes in HIMU lavas reveal Archean crust in their mantle source. *Mineralogical Magazine*. 77(5):805.
19. TRAIL, Dustin, WATSON, Bruce E., and TAILBY, Nicholas D. (2011): The oxidation state of Hadean magmas and implications for early Earth's atmosphere. *Nature*. 480:79-82.
20. KENDRICK, Mark A., SCAMBELLURI, Marco, HONDA, Masahiko, *et al.* (2011): High abundances of noble gas and chlorine delivered to the mantle by serpentinite subduction. *Nature Geoscience*. 4:807-812.
21. ANZELLINI, Simone, DEWAELE, Agnes, MEZOUAR, Mohamed, *et al.* (2013): Melting of Iron at Earth's Inner Core Boundary Based on Fast X-ray Diffraction. *Science*. 340(6131):464-466.

REFERENCES

- BABYCH, Bogdan, HARTLEY, Anthony, KAGEURA, Kyo, *et al* (2012): MNH-TT: a collaborative platform for translator training. *Proceedings of Translating and the Computer*. 34:1-17.
- BAKER, Mona (1998): Réexplorer la langue de la traduction: une approche par corpus. *Meta*. 43(4):480-485.
- BAKER, Mona (1999): The Role of Corpora in Investigating the Linguistic Behaviour of Professional Translators. *International Journal of Corpus Linguistics*. 4(2):281-298.
- BONWELL, Charles and EISON, James (1991): *Active Learning: Creating Excitement in the Classroom*. ASHE-ERIC Higher Education Report. No. 1. Washington, D.C.: School of Education and Human Development (The George Washington University). Visited 2 April 2018, <<https://files.eric.ed.gov/fulltext/ED336049.pdf>>.
- BOWKER, Lynne and BENNISON, Peter (2003): Student Translation Archive and Student Translation Tracking System. Design, Development and Application. In: Federico ZANETTIN, Silvia BERNARDINI, and Dominic STEWART, eds. *Corpora in Translator Education*. Manchester: St. Jerome, 103-118.
- CASTAGNOLI, Sara, CIOBANU, Dragos, KÜBLER, Natalie, *et al* (2011): Designing a Learner Translator Corpus for Training Purposes. In: Natalie KÜBLER, ed. *Corpora, Language, Teaching, and Resources: From Theory to Practice*. Bern: Peter Lang, 221-248.
- FLEURY, Serge and ZIMINA, Maria (2014): Trameur: A Framework for Annotated Text Corpora Exploration. In: Lamia TOUNSI and Rafal RAK, eds. *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: System Demonstrations*. (COLING 2014: 25th International Conference on Computational Linguistics, Dublin, 23-29 August 2014). Dublin/Stroudsburg: Dublin City University/Association for Computational Linguistics, 57-61.
- FRANKENBERG-GARCIA, Ana (2009): Are translations longer than source texts? A corpus-based study of explicitation. In: Allison BEEBY, Patricia RODRÍGUEZ-INÉS, Pilar SÁNCHEZ-GIJÓN, eds. *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, 47-58.
- FROELIGER, Nicolas (2013): *Les Noces de l'analogique et du numérique – De la traduction pragmatique*. Paris: Les Belles lettres.

- GÖPFERICH, Susanne (2009): Towards a Model of Translation Competence and its Acquisition: The Longitudinal Study of TransComp. In: Sussanne GÖPFERICH, Arnt Lykke JAKOBSEN, and Inger M. MEES, eds. *Behind the Mind: Methods, Models, and Results in Translation Process Research*. Frederiksberg: Samfundslitteratur Press, 12-39.
- KÜBLER, Natalie (2003): Corpora and LSP translation. In: Federico ZANETTIN, Silvia BERNARDINI, and Dominic STEWART, eds. *Corpora in Translator Education*. Manchester: St. Jerome, 25-42.
- KÜBLER, Natalie, MESTIVIER, Alexandra, PECMAN, Mojca, et al (2016): Exploitation quantitative de corpus de traductions annotés selon la typologie d'erreurs pour améliorer les méthodes d'enseignement de la traduction spécialisée. In: Damon MAYAFFRE, Céline POUDAT, Laurent VANNI, et al., eds. *Actes des 13^{ème} Journées internationales d'Analyse statistique des Données Textuelles - International Conference on Statistical Analysis of Textual Data (JADT)*. (JADT2016: 13^{ème} Journées internationales d'Analyse statistique des Données Textuelles, Nice, 7-10 June 2016). Nice: Université Nice Sophia Antipolis, 731-742.
- LADMIRAL, Jean-René (1994): *Traduire: théorèmes pour la traduction*. Paris: Payot.
- LADMIRAL, Jean-René (2014): *Sourcier ou cibliste*. Paris: Les Belles lettres.
- LOOCK, Rudy, MARIAULE, Michaël, and OSTER, Corinne (2013): Traductologie de corpus et qualité: étude de cas. In: *Tralogy II*. (Deuxième colloque Tralogy: anticiper les technologies pour la traduction, Paris, 17-18 January 2013). Vandœuvre-lès-Nancy: Institut de l'Information Scientifique et Technique. Visited 9 February 2018, <<http://odel.irevues.inist.fr/tralogy/index.php?id=243>>.
- MAURANEN, Anna (2007): Universal Tendencies in Translation. In: Margaret ROGERS and Gunilla ANDERMAN, eds. *Incorporating Corpora: The Linguist and The Translator*. Clevedon/Tonawanda: Multilingual Matters, 32-48.
- MESTIVIER-VOLANSCHI, Alexandra (2015): Productivity and Diachronic Evolution of Adjectival and Participial Compound Pre-modifiers in English for Specific Purposes. *Fachsprache*. 37(1-2):2-23.
- OLOHAN, Maeve (2004): *Introducing Corpora in Translation Studies*. London/New York: Routledge.
- OLOHAN, Maeve and BAKER, Mona (2000): Reporting *that* in Translated English: Evidence for Subconscious Processes of Explication? *Across Languages and Cultures*. 1(2):141-158.
- PECMAN, Mojca (2014): Variation as a cognitive device: how scientists construct knowledge through term formation. *Terminology*. 20(1):1-24.
- PECMAN, Mojca and KÜBLER, Natalie (2011): ARTES: an online lexical database for research and teaching in specialised translation and communication. In: Benoît SAGOT, eds. *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011*. (WoLeR: International Workshop on Lexical Resources, Ljubljana, August 1-5 2011). Ljubljana: European Summer School in Logic, Language and Information, 86-93.
- PACTE Group (2000): Acquiring Translation Competence: Hypotheses and Methodological Problems in a Research Project. In: Allison BEEBY, Doris ENSINGER, and Marisa PRESAS, eds. *Investigating Translation*. Amsterdam/Philadelphia: John Benjamins, 99-106.
- PEARSON, Jennifer (2003): Using Parallel Texts in the Translator Training Environment. In: Federico ZANETTIN, Silvia BERNARDINI, and Dominic STEWART, eds. *Corpora in Translator Education*. Manchester: St. Jerome, 15-24.
- PUURTINEN, Tiina (2003): Genre-specific Features of Translationese? Linguistic differences between Translated and Non-translated Finnish Children's Literature. *Literary and Linguistic Computing*. 18(4):389-406.
- ROGERS, Margaret (2015): *Specialised Translation. Shedding the 'Non-literary' Tag*. Basingstoke/New York: Palgrave Macmillan.
- SWALES, John M. (1990): *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.