ARTÍCULO DE INVESTIGACIÓN ORIGINAL

ENHANCING TRANSLATION COMPETENCE THROUGH AN ENGLISH-SPANISH SENTENCE-ALIGNED PARALLEL CORPUS OF SCIENCE AND TECHNOLOGY TEXTS

DESARROLLO DE LA COMPETENCIA TRADUCTORA A TRAVÉS DE UN CORPUS PARALELO ALINEADO DE TEXTOS CIENTÍFICO-TÉCNICOS

Lisset Rodríguez-Águila, [1]   Humberto Miñoso-Machado, [2] Juana Idania Pérez-Morales [3]

[1]5th Year Student of the Degree Course: English Language with Second Foreign Language (French) at Universidad Central "Marta Abreu" de Las Villas, Cuba. Correo: lissetra223@gmail.com ORCID ID: https://orcid.org/0000-0002-6877-7313. [2] M.Sc. Faculty of English Language, Associate Professor Universidad Central "Marta Abreu" de Las Villas, Cuba. Participates in the research projects: "Development of the communicative and intercultural competences for a sustainable human development" (institutional with national scope) and "Accelerated learning of the English Language for Professionals" (institutional scope). Correo: humbertomm@uclv.edu.cu ORCID ID: https://orcid.org/0000-0001-7716-1572. [3] Full Professor of the English Language Department at Universidad Central "Marta Abreu" de Las Villas. PhD in Pedagogical Sciences. Participates in the research projects: "Cuban and Latin-American identity and studies on Humanities in the central region of Cuba" (national scope), "Learning Assessment for the development of English communicative language skills of students at the Universidad Central "Marta Abreu" de Las Villas (institutional with national scope) and "Development of the communicative and intercultural competences for a sustainable human development" (institutional with national scope). Correo: juanap@uclv.edu.cu ORCID ID: https://orcid.org/0000-0002-3121-9254

_____

_____

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

291

Abstract

This article describes the use of an English-Spanish Parallel Corpus of Science and Technology Texts which pursues the objective of enhancing the translation competence of the fourth-year students of the degree course English Language with Second Foreign Language (French), at Universidad Central "Marta Abreu" de Las Villas. The research adopted a mixed-method approach design for collecting, analyzing, and interpreting the data related to the translation of science and technology texts, and to diagnose the needs of the students in that respect. The design of the Corpus followed the guidelines of corpus linguistics as a methodology. A corpus of 50 science and technology texts in English, and their translations into Spanish, on the domains of medicine, physics, chemistry and artificial intelligence, was compiled and processed using Corpus processing tools such as *OmegaT* aligner and *Xbench* software. It provides a wide range of lexical, syntactical and textual examples in English with their equivalents in Spanish, in order to help students integrate the necessary knowledge and skills to translate those texts, through a process-oriented approach to translation. The corpus derived a word bank of science and technology terms and a proposal of corpus-based exercises to be used in teaching translation.

Keywords: parallel corpus, science and technology texts, translation competence

RESUMEN

Este artículo describe el uso de un Corpus Paralelo de Textos Científico-Técnicos para mejorar la competencia traductora de los estudiantes de cuarto año de la carrera de Lengua Inglesa con Segunda Lengua Extranjera (Francés), en la Universidad Central "Marta Abreu" de Las Villas. La investigación adoptó un enfoque mixto para la recopilación, análisis e interpretación de datos relacionados con la traducción de textos científico-técnicos, y para diagnosticar las necesidades de los estudiantes al respecto. El diseño del corpus siguió las directrices de la lingüística de corpus como metodología. Se compiló un corpus de 50 textos de ciencia y tecnología en inglés, y sus traducciones al español, en los dominios de medicina, física, química e

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb.,* 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

292

inteligencia artificial, y se procesó utilizando herramientas de trabajo con corpus como el alineador *OmegaT* y el software *Xbench*. Proporciona una amplia gama de ejemplos léxicos, sintácticos y textuales en inglés con sus equivalentes en español, para ayudar a los estudiantes a integrar los conocimientos y habilidades necesarios mediante un enfoque hacia el proceso de la traducción. Del corpus se deriva un banco de términos científico-técnicos y una propuesta de ejercicios basados en el corpus para ser utilizados en la enseñanza de la traducción.

Palabras clave: competencia traductora; corpus paralelo; textos científico-técnicos

INTRODUCTION

Currently, learners of English as a Foreign Language in higher education institutions face a plethora of challenges. Many of them are related to the development of specific competences that are of paramount importance for their professional performance as translators and interpreters. In the context of the degree course English Language with Second Foreign Language at Universidad Central "Marta Abreu" de Las Villas, the development of translation competence has become a priority for both, professors and learners, particularly in the field of specialized translation, in which the topics pose a double degree of linguistic difficulty laying in the knowledge of the specific terminology, that needs to be overcome.

The proposal of an English-Spanish Sentence-Aligned Parallel Corpus of Science and Technology Texts will help the students of the degree course English Language with Second Foreign Language (French) to acquire and develop their translation competence as undergraduates, and hopefully, in their working environment as translators. According to experts, using parallel corpora as an aiding tool in the translation process, provides a broad range of vocabulary and translation solutions, which is beneficial to systematize lexical knowledge and translation strategies. Considering the importance of enhancing translation competence in the students of the aforementioned degree course, the proposal could be viewed as an innovative and far-reaching alternative.

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

293

No one knows for certain how many scientific journals there are, but several estimates point to around 30,000 with close to two million articles published each year. (Philip G Altbach and Hans de Wit, 2018) With more and more works and papers on science and technology published in English, and their level of lexical complexity, the need to enhance the competence of the translators is imperious, since they are the bridge between the scientists and the readers. Hence, the relevance of this research work.

Statement of the problem

The difficulties evidenced in this research context are related with the lack of aiding tools to raise fourth-year English Language students' translation competence in the classroom, particularly when dealing with topics such as science and technology, which present a higher level of difficulty in translation due to their specialized terminology, academic language and formal style, and even due to the complexity of topics, which are not of students' common knowledge. It is also noticeable an absence of systematic study and practice based upon the analysis of already-made translations, which is an approach to translation

teaching that could help students to learn and apply a broad range of solutions to difficulties in the process of translation, either at the lexical or syntactic level, besides visualizing the main stylistic and linguistic features of the texts studied.

The problem evidenced leads the authors of this paper to formulate the following **research question**: How could the translation competence of the fourth-year English Language students be enhanced, in order to prepare them for accurately translating science and technology texts?

Consequently, the **overall aim** of this research work is:

To describe the use of an English-Spanish Sentence-Aligned Parallel Corpus of Science and Technology Texts for enhancing fourth-year English Language students' translation competence through a process-oriented approach to translation.

This proposal entails to incorporate the use of a parallel corpus as part of the acquisition of the translation competence, being a useful tool to provide solutions for translating difficult lexical items, for comparing the source and target texts and see the solutions

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb.,* 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

294

given by other translators to lexical or phrasal items that are difficult for Spanish speakers, or that are specialized scientific and technical terminology. This proposal pursues to be an aid that serves to integrate the necessary skills and knowledge to translate scientific and technical texts, referring to the systematization of specialized terminology, and to incorporate to the learning process a wide variety of examples that provide students with various alternatives to solve particular translation difficulties, all of which will lead them to enhance their translation competence.

THEORETICAL FRAMEWORK

Defining Translation Competence

In the decades of the 80s and 90s of the 20th century, many linguists and translation theorists attempted to define translation competence and its underlying sub-competences. These experts laid the basis for more recent definitions that clearly explain what translation competence is, and what specific sub-competences are integrated into the main concept.

For instance, Schäffner (2000) and Šeböková (2010) cited in Abouzar Oraki & Gholam-Reza Tajvidi (2020),

identified the translation competence as a set of different sub-competences required for translating. Both of them identified six main translation sub-competences. The first author proposed the following ones: 1) Linguistic competence of the languages concerned, 2) Cultural competence, 3) Textual competence, 4) Domain/subject specific competence, 5) (Re)search competence and 6) Transfer competence.

In turn, the translation sub-competences proposed by the second author are more complete. They include: 1) Core Translation Competence (which integrated all the other sub-competences), 2) Linguistic sub-competence, 3) World-subject competence, 4) Research competence, 5) Tools competence and 6) Cultural competence.

Moreover, the PACTE group (*Proceso de Adquisición de la Competencia Traductora y Evaluación*, by their acronym in Spanish) of the *Departament de Traducción i d'Interpretació de la Universitat Autonóma de Barcelona*, a pioneer in the study of translation competence with empiric and experimental studies also define the concept of translation

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb.,* 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

295

competence. This group of scholars define the translation competence as:

The underlying knowledge system needed to translate". It has four distinctive characteristics: it is expert knowledge and not possessed by all bilinguals, it is basically procedural knowledge, it is made up of various interrelated sub-competences, and the strategic component is very important, as it is in all procedural knowledge. (PACTE, 2003, 43-66)

They consider that the overall aim of translation teaching is the development of the translation competence, which is enhanced through training and systematic study since it is considered basically a procedural knowledge. The various interrelated sub-competences they proposed are:

1. *The bilingual sub-competence:* Predominantly procedural knowledge needed to communicate in two languages. It includes the specific feature of interference control when alternating between the two languages.

It is made up of pragmatic, socio-linguistic, textual, grammatical and lexical knowledge in the two languages.

2. *Extra-linguistic sub-competence:* Predominantly declarative knowledge, both explicit and implicit, about the world in general and specific areas. It includes (1) bicultural knowledge (about the source and target culture); (2) encyclopedic knowledge (about the world in general); and (3) subject knowledge (in special areas).

3. *Knowledge about translation sub-competence:* Predominantly declarative knowledge, both implicit and explicit, about what translation is and aspects of the profession. It includes (1) knowledge about how translation functions: type of translation units, processes required, methods and procedures used (strategies and techniques), and types of problems; (2) knowledge related to professional translation practice: knowledge of the work market (different types of briefs, clients and audiences, etc.).

4. *Instrumental sub-competence:* Predominantly procedural knowledge related to the use of documentation sources and information and communication technologies applied to

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

296

translation: dictionaries of all kinds, encyclopedias, grammars, style books, parallel texts, electronic corpora, searchers, etc.

5. Strategic sub-competence: Procedural knowledge to guarantee the efficiency of the translation process and solve the problems encountered. This is an essential sub-competence that affect all the others and cause inter-relations amongst them because it controls the translation process. Its functions are (1) cognitive components such as memory, perception, attention and emotion; (2) attitudinal aspects such as intellectual curiosity, perseverance, rigor, critical spirit, knowledge of and confidence in one's own abilities, the ability to measure one's own abilities, motivations, etc.; (3) abilities such as creativity, logical reasoning, analysis and synthesis, etc.

Translation competence is, in short, expert knowledge (declarative, procedural, and attitudinal) that requires analytical and intuitive thinking, made up of various interrelated sub-competences which requires the specialization of the individuals. The acquisition of translation competence occurs mainly through a process-oriented approach, which focuses on the analysis of the steps taken by the translator when translating a text, the possible solutions, the acquisition of principles and techniques, etc.; it focuses on the process rather than on the result of the translation exercise. (Gile, 2018)

Defining corpus linguistics. Corpus typology.

What is corpus linguistics and what is a corpus? According to professor Sara Cushing (2017), *Corpus linguistics* is a research methodology used for the study of linguistic phenomena through large collections of machine-readable texts, oral or written, called *corpora*. In other words, it is a research methodology based on empirical evidence about how language is actually used.

Now, corpora can be classified into various categories, considering the source of the content, metadata, number of languages, objective, etc. According to the EAGLES (Expert Advisory Group on Language Engineering Standards) text typology, the most general and first distinction made between corpora occurs between Monolingual and Multilingual (including Bilingual) corpora. Multilingual

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903*
Email: pedagogiasociedad@uniss.edu.cu

297

(bilingual) corpora can be sub classified into Comparable corpora (corpora compiled using similar design criteria but which are not translations) and Parallel, or Translation Corpora (Zanettin, 2012).

Having adddressed individually the main theoretical concepts related to translation competence and corpus linguistics, the following question is raised: How can corpus linguistics specifically contribute to the enhancement of translation competence and each underlying translation sub-competence?

Using a parallel corpus in the translation classroom for the enhancement of translation sub-competences

Within the whole typology of corpora with specific uses in linguistic studies, the interest of this research focuses on a specific type of corpus called parallel corpus. A *parallel corpus* comprises texts in one language aligned with their translations in another. Both languages need to be aligned, i.e. identifying corresponding segments, usually sentences or paragraphs, which need to be matched through a corpus tool. (Zanettin, 2012)

Parallel corpora are considered one of the most helpful tools in the enhancement of the translation competence. In this vein, Monserrat Bermúdez Bausela (2016) strongly affirms that corpora, particularly specialized corpora, are an inestimable source not only for terminology and phraseology extraction, but also for studying the textual conventions that characterize and define specific genres in the translation languages. In her paper "*The importance of corpora in translation studies: a practical case*", she highlights the contribution of corpora to the study of a specialized language from the translator's point of view, which specifically enables the enhancement of the bilingual translation sub-competence.

> *From the observation of specialized corpora, it is possible to identify specific patterns, phraseology, terminological variants, and the frequency of conceptually relevant words, cohesive features and so forth. The access to this information will allow the translator to*

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903*
Email: pedagogiasociedad@uniss.edu.cu

298

produce   quality   text.
(Bausela, 2016)

One of the main advantages of translation teaching with corpora over traditional translation teaching is generally considered to be the fact that the presence of corpora reduces the role of the teacher's intuition in the translation classroom and, at the same, time assigns more importance to the students and their documentation skills (Rodríguez Inés, 2008) cited in (Krüger, 2020). By providing alternative sources of authority as well as a set of authentic data, corpora can also shift the role of the teacher from that of the principal information provider to that of an information facilitator (Rodríguez Inés, 2008) cited in (Krüger, 2020).

According to Professor Patricia Rodríguez-Inés (2008), an electronic parallel corpus:

1. Can be built in less time than a conventional printed corpus. (which specifically contributes to the enhancement of the instrumental sub-competence)
2. Can be reutilized (which specifically contributes to the enhancement of the instrumental and knowledge about translation sub-competences).
3. Allows to carry out quick and systematic search of information in a wider and more varied range of texts. (instrumental, strategic and extra linguistic sub-competences)
4. Offers a great number of lexical examples in the source and target languages in context (which specifically contributes to the enhancement of the bilingual sub-competence).
5. Allows translators to go directly to those sections in the text where there are key words with the assistance of a corpus analysis tool (which specifically contributes to the enhancement of the instrumental sub-competence).
6. Facilitates the detection of information as collocational patterns (bilingual-sub-competence).
7. Facilitates the detection of frequency, which allows to determine the relevance of the patterns found (knowledge about translation sub-competence).
8. Facilitates and accelerates the retrieval and analysis of concrete

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903*
Email: pedagogiasociedad@uniss.edu.cu

299

information (textual sub-competence).

9. Offers certitude at the time of making strategic decisions such as lexical choices (which specifically contributes to the enhancement of the strategic sub-competence).

10. Provides a wide range of examples on how other translators dealt with specific translation problems, and this knowledge can be incorporated by students in order to solve similar translation problems (strategic and knowledge about translation sub-competences).

11. Provides greater opportunities for casual learning (which specifically contributes to the enhancement of the extra-linguistic sub-competence). (Rodríguez-Inés, 2008)

All in all, experts coincide about the positive outcome of using corpora, specifically parallel corpora, to aid the translation process and the enhancement of the different translation sub-competences (Frérot, 2016). So far, having established the theoretical foundations related to the use of parallel corpora as an aiding tool in the translation process and the enhancement of the translation competence, it is pertinent to move forward by enunciating the methodology and practical steps involved in the construction of the parallel corpus proposed in the present research work.

METHODOLOGY

Methodology for building an English-Spanish Sentence-Aligned Parallel Corpus of Science and Technology Texts

The design of the present research adopted a mixed-method approach. This approach analyzes the collected data quantitatively and qualitatively interpret the research results.

Sixteen fourth-year students of the English Language with Second Foreign Language (French) degree course at Universidad Central "Marta Abreu" de Las Villas were surveyed in order to diagnose the current situation concerning the students' perceptions on the translation process when translating science and technology texts, as well as their considerations regarding the needs and deficiencies in the process. These students were chosen because they are the ones who translate science

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

300

and technology texts in their translation lessons.

Moreover, a collection of 50 science and technology texts were chosen to build the English-Spanish Sentence-Aligned Parallel Corpus proposed in this paper. The corpus dataset comprised texts in English and their translations in Spanish, consistent with the topics dealt in the translation lessons. The texts were retrieved from reliable sources and some of them were provided by the professor of the subject.

Diagnosing the perception of students on the subject Translation of Science and Technology Texts

A survey was applied to 16 fourth-year English Language with Second Foreign Language (French) students in the academic year 2019-2020, with the purpose of diagnosing the current situation concerning the students' translation competence, specifically in the Translation of Science and Technology Texts.

The criteria for the selection of the sample was that the subject Translation of Science and Technology Texts is taught in this specific year of the aforementioned degree course. The survey also introduced the possibility of using a bilingual parallel corpus as an aiding tool for the practice of translation, and to help students enhance their translation competence in the subject Translation of Science and Technology texts, which, as stated before, is very complex.

Results of the survey: The majority of students (90 percent) perceived some difficulties hindering the translation process:

1. Insufficient offline bilingual reference sources for translating in the lessons.
2. Insufficient time for translating in class.
3. Some students claimed for more tutoring hours with the professors.
4. Most of the translation is done through independent study.
5. Insufficient background information on the topics of science and technology before facing the translation of a text dealing with these topics.

In an open question about the importance of using a parallel corpus like the one proposed in this research,

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

301

many students agreed that a corpus could be a useful tool in the translation process, given the absence of systematic study and practice based upon the analysis of already-made translations to strengthen the translation competence of students.

Some of the students´ opinions are shown below to support the aforesaid statement:

Student 3: *"It would be very useful as a tool, not only for counting on a reliable source for term translation, but also to see them in context and with a variety of examples, collocations and alternatives of specific professionalisms".*

Student 7*: "It would be very helpful. We could see the use of terms in context and finish our translation quicker and with more quality".*

Student 8: *"Not only it would be helpful in terms of contrasting and finding new solutions to the translation problems we face in the translation of scientific and technological texts, but it would help us to practice, to enhance not only the result but the process of translation, to learn and systematize new translation strategies".*

Table 1. Results of the survey. Quantitative analysis

Students' perceptions concerning the subject Translation of Science and Technology Texts			
Availability of bilingual reference sources	Enough	Little	Insufficient
	7.69%	46.15%	38.46%
Prior knowledge of background information, vocabulary and possible solutions to translation difficulties	Enough	Little	Insufficient
	23.07%	23.07%	53.84%
Dependence on online translators	Very little dependent	More or less dependent	Very dependent
	23.07%	30.76%	46.15%
Using only online	Yes	No	

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

302

translators is enough to improve translation competence	7.69%	92.3%	-
Students who have used a bilingual corpus as an aiding tool to translate	-	100%	-

Source: Elaborated by the author (2020)

The results derived from the quantitative and qualitative analysis of the survey, led the authors to conclude that there is a gap in the development and enhancement of the translation competence of the fourth-year students, particularly referring to science and technology texts. Overall, students do not feel prepared before and while translating, in addition to the insufficient systematization and integration of the acquired knowledge, which gives learners the assurance of being prepared to face any scientific or technical translation.

Steps for building the English-Spanish Sentence-Aligned Parallel Corpus of Science and Technology Corpus design criteria

In the design of a corpus, it is important to consider and recall constantly that a corpus is a subset of an electronic library, built accordingly to explicit design criteria for a specific purpose. Having taken that into consideration, it is also necessary to establish a planning process and to use it as a guide in the construction of the corpus. According to specialists' criteria, such planning should include the following stages:

i. Selection of sources

ii. Data capture and encoding

iii. Corpus processing (Sue Atkins, Jeremy Clear and Nicholas Ostler, 2018)

For building the corpus, it was necessary to use Corpus processing tools such as the *OmegaT* aligner (Keith Godfrey & Aaron Madlon-Kay, 2000), and the *Xbench* tool (ApSIC, 2017). All the files were converted to

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

303

.txt format and subsequently encoded in UTF8 format. Once all the files were encoded in the required formats, they were uploaded to the *OmegaT* tool to be aligned at a sentence level. Afterwards, they were uploaded to *Xbench* and converted into a translation memory, in order to be stored and allowing the users to carry out word search. Finally, *AntConc* tool (Lawrence, 2014) was used for analyzing the most common translation difficulties and specialized terminology, in order to organize them into a word bank.

Selection of sources

The texts included in the Sentence-Aligned English-Spanish Parallel Corpus of Science and Technology Texts are a very specific sample chosen in accordance with the subjects dealt with in the translation discipline of the English Language with Second Foreign Language (French) degree course.

So far, the collection of text comprises 50 texts and their respective translations, retrieved from reliable sources such as the WHO (World Health Organization) official website, the Nobel Prizes official website, and the Encarta Encyclopedia Student Premium 2009. Some of these texts have been used by fourth-year students and professors of the English Language with Second Foreign Language (French) degree course in their translation lessons, and have been already corrected in group discussions with the help of the professor.

In the first semester of 4th year, students work with science and technology texts in the domains of: (1) medicine and (2) scientific discoveries on physics, chemistry and artificial intelligence. These domains are divided into the specific topics of: communicable and non-communicable diseases, artificial intelligence models and Nobel Prizes. Due to the recent impact of the Covid-19 pandemic, it has been considered relevant to include a number of texts dealing with this topic in the corpus, which belongs to the domain of medicine.

The characteristics of the selected corpus follow the compilation criteria exposed by Vargas (2006). These characteristics are shown in Table 2 below:

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb.,* 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

304

Table 2. Characteristics of the selected corpus

Size of the corpus	125 680 words
Transmission mode	Written transformed into .txt format
Level of specialization	Specialized texts of specific domain of medicine, physics, chemistry and artificial intelligence
Type of texts	Published research articles
Authorship	Varies from science specialist explaining a discovery to journalist reporting a specific scientific milestone
Language of the corpus	Texts originally written in English
Date of publication	Texts published from 2009 to 2020

Vargas Sierra, Ch. (2006). Diseño de un corpus especializado con fines terminográficos: el corpus de la piedra natural. *Debate Terminológico*, 2 (7/2006). París: RITERM (Red Iberoamericana de Terminología).

Data capture and encoding

An aligned corpus results from a process called alignment, which can be defined as "finding correspondences, in bilingual parallel corpora, between textual segments that are translation equivalents" (Kraif, 2018). In the case of the corpus in question, the alignment was made at a sentence level because it is more user-friendly and more comfortable for the automatic work.

The tool used to align the proposed corpus was *OmegaT*, a free and open source multiplatform Computer Assisted Translation tool with a very complete set of utilities to edit translated texts such as fuzzy matching, translation memory, keyword search, glossaries, and translation leveraging into updated projects. This tool has a very reliable aligner. As has been referred above, in order to work with a corpus analysis tool or a translation-aiding tool, it is necessary that the texts are in a machine-readable format. In this case, texts in both languages were converted to txt. format and subsequently to UTF8. format, which made it possible to upload and align the texts through the *OmegaT* aligner, as shown in **Screenshot 1**.

Pedagogía y Sociedad | Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

305

Screenshot 1. Automatic alignment and manual edition of the corpus

Corpus processing

The corpus processor tool used to store the English-Spanish Sentence Aligned Parallel Corpus is called *Xbench,* a user-friendly and accessible tool. It is available in free versions on the Internet. Once the parallel corpus has been uploaded as a new project in *Xbench*, this tool allows users to carry out word and phrase search in both, the source and the target language, as shown in **Screenshot 2**. In the Science and Technology Sentence-Aligned English-Spanish Parallel Corpus, the searched words in the source language and their equivalents appear aligned and in context, thanks to the alignment process previously carried out. The English-Spanish Sentence-Aligned Parallel Corpus was saved in an *Xbench* file, as a translation memory, and from now on will be called *ParaCorp2020.*

Screenshot 2. Word search in *ParaCorp2020* (using *Xbench* software)

RESULTS AND DISCUSSION

Proposal of a word bank and exercises to use *ParaCorp2020* in the translation lessons

Derived from an analysis of the corpus carried out with *AntConc* corpus tool, the most relevant scientific and technical (ST) terminology was extracted from the texts of the corpus, in order to compile a word bank. The technical-scientific terms extracted were organized in a word bank of 1872 words, which covers the most relevant vocabulary related to the topics dealt with in the translation lessons. It is expected that this word bank will be a powerful aiding material for the acquisition and systematization of specialized terminology, necessary for translating the texts in class and for the future professional development of the students as translators.

In addition to the word bank of technical-scientific terminology as a supporting material and pedagogic outcome of *ParaCorp2020,* the author of this research also proposes a second pedagogic use for the corpus in the translation lessons. This proposal consists of a sample of exercises which have the purpose of motivating the

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb.,* 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

306

students and introducing the use of *ParaCorp2020* in the translation class as a didactic tool, in order to help students to enhance their translation competence.

The objectives of this proposal of exercises are:

i. To facilitate the acquisition of knowledge and skills through the use of *ParaCorp2020* for translation.

ii. To enable the acquisition of knowledge and skills necessary to translate about ST texts.

iii. To provide a variety of contextualized examples of equivalents in both languages, so that students can incorporate the ST vocabulary that appears in the texts and that sometimes is not found in the dictionary.

iv. Provide a range of solutions to different translation problems due to differences in grammatical and syntactic structures of both languages.

v. To provide didactic activities for practicing translation, which do not only consist in "reading and

translating", but also intend that students can make generalizations, discover principles and necessary skills when translating.

Each exercise sheet will work with a different text. They will serve as a lesson´s guidelines and will have an *introduction*, a *body* and *conclusions* as an ordinary lesson. The *introduction* comprises a warm-up section with general familiarization questions so that the students can have a preliminary idea of the topic they will be dealing with during the lesson, as well as to check the background knowledge they have about it. After the warm-up, it is important to state which field and specific topic will be treated in the lesson, and to establish its objectives.

The *body* of the exercise sheet will comprise the presentation of the English-Spanish Sentence-Aligned Parallel Corpus of Science and Technology Texts (at least in a first lesson), and the corpus-based translation exercises to be completed by the students. These exercises are interactive, so that the teacher and the students can search together for possible alternatives, and strengthen

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903*
Email: pedagogiasociedad@uniss.edu.cu

307

the translation process. The techniques used in the lesson should not be based on the prevalence of the teacher´s solutions over the students´, but on a mutual collaboration and dialogue as to find the best solutions and most importantly, that the students can systematize new terms, translation principles and solutions. They need to focus on the process and the acquisition of techniques, without neglecting the quality of the results. The acquisition and development of translation competence through the continuous practice, will contribute to the improvement of the translations. Finally, in the *conclusions,* the teacher and students will review all the translation difficulties, new techniques, and new terminology learned during the lesson in order to systematize and fix the knowledge acquired.

Table 3 Layout of the corpus-based exercises

Introduction	**Warm-up.** State field, topic and objectives of the lesson
Body	Presentation of *ParaCorp2020* and the text to translate in class*.* Complete 1-10 corpus-based translation exercises. Carry out interactive revision of the translation, as well as an analysis of the difficulties and its possible solutions through the search of alternatives in the corpus and interactive debate in class.
Conclusions	Summarize the new terminology and techniques learned in class. Evacuate possible doubts. Ask students for feedback about the translation exercises and the use of *ParaCorp2020.*

Elaborated by the author (2020)

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb.,* 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

308

CONCLUSIONS

The integration of language disciplines has become vital to foster linguistic competences needed for the successful professional development of translators and interpreters. Many experts highlight the benefits of the integration of corpus linguistics and translation studies, and its practical applications to the training of novice translators.

The needs analysis revealed the lack of aiding tools to raise students' translation competence when dealing with the translation of science and technology texts, the lack of terminological knowledge and strategic skills to translate led to the proposal of an English-Spanish Sentence-Aligned Parallel Corpus of Science and Technology Texts (***ParaCorp2020***) as an authentic and innovative idea to aid the enhancement of the translation competence.

 The corpus derived a word bank of science and technology texts and a proposal of corpus-based exercises which will also contribute to the purpose of the proposal.

REFERENCES

Abouzar, Oraki & Gholam-Reza, Tajvidi. (2020). Training Translators and Interpreters: The Need for a Competence-Based Approach in Designing University Curricula. *Iranian Journal of English for Academic Purposes*, (1).

ApSIC Team. (2017). *Easy Quality Assurance and Terminology.* Retrieved from http://www.xbench.net/

Bausela, M. B., 2016. The importance of corpora in translation studies: a practical case. *Dublin, Research-publishing.net.* ERIC, Institute of Education Sciences. Retrieved from: https://files.eric.ed.gov/fulltext/ED565828.pdf.

Cushing, S. (2017). Corpus linguistics in language testing research. *Sage Journals, 44*, 1.

Frérot, C. (2016). *Corpora and Corpus Technology for Translation Purposes in Professional and academic Environments. Major Achievements and New Perspectives.* Universidad de Grenoble, France.

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb., 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

309

Gile, D. (2018). The Process-Oriented Approach in Translation Teaching. In C. D. Lindegaard, *Teaching Translation and Interpreting 2: Insights, aims, visions.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kraif, O. (2018). *From Translational Data to Contrastive Knowledge: Using bi-texts for Bilingual Lexicon Extractions. International Journal of Corpus Linguistics.*

Krüger, R. (2020). *Working with corpora in the translation classroom. Studies in Second Language Learning and Teaching.*

Lawrence, A. (2014). *AntConc* (Version 3.4.3w). [Computer Software]. Waseda University, Tokyo, Japan. Retrieved from http://www.antlab.sci.waseda.ac.jp/

PACTE (2003). *Building a Translation Competence Model. Triangulating Translation: Perspectives in process oriented research.* Amsterdam, Paises Bajos.

Rodríguez-Inés, P. (2008). *Uso de Corpus Electrónicos en la Formación de Traductores (inglés-español-inglés).* (doctoral thesis). Barcelona, España.

Šeböková, S. (2010). *Comparing translation competence* (master's thesis). Masaryk University, Brno, Czech Republic.

Sue Atkins, J. C. & N. Ostler (2018). Corpus Design Criteria. *Literary and Linguistic Computing* , 3.

Zanettin, F. (2012). *Translation-Driven Corpora: Corpus Resources for Descriptive and Applied Translation Studies.* Manchester: St. Jerome Publishing, 10.

Pedagogía y Sociedad | *Vol. 24, no. 62, nov.-feb.,* 2021, ISSN 1608-3784. RNPS: 1903
Email: pedagogiasociedad@uniss.edu.cu

310