

Do similar contexts of production entail similar textual features? A corpus-based comparison of the simplification-related features in translation and Academic Writing

Virginia Mattioli
Investigadora
virginiamattioli@gmail.com

Enviado: 19.10.2021

Revisado: 3.11.2022

Aceptado: 22.12.2022

Abstract: This paper explores the relationship between the context of production of texts and their specific features. With this aim, the study compares the features related to syntactic and stylistic simplification between two textual practices: translation and Academic Writing. Actually, both practices are produced in common contexts with similar purposes. Further, the exam of simplification-related features in different sets of texts translated into English from different languages also allow for further verifying the translation universals hypothesis of simplification. For the analysis, an archive of four comparable corpora of English texts is compiled, representing original and translated versions of academic and non-academic texts, respectively. Once determined concrete indicators of simplification (lexical variety, lexical density, mean sentence length, presence of subordination and of non-finite sentences), a corpus-based methodology is used to identify each one of them in each set of texts, according to the hypothesis that both practices present the same simplification-related features at the same extent. The comparison of the results refutes both hypotheses, showing that each practice presents different simplification-related features and, in all the analyzed corpora, originals result to be simpler than translations. Moreover, the distribution of the examined indicators across the four corpora seems to follow a regular pattern suggesting an interesting point of departure for further studies.

Key words: Corpus-based translation studies; Translation universals; Simplification; Academic writing; Context of textual production.

¿Implican contextos similares de producción características textuales similares? Comparación basada en corpus de las características relacionadas con la simplificación en la traducción y la redacción académica

Resumen: Este artículo pretende explorar la existencia de alguna relación entre el contexto de producción de los textos y sus rasgos específicos. En concreto, se comparan los rasgos relacionados con la simplificación sintáctica y estilística entre dos prácticas textuales: la traducción y el discurso académico. Efectivamente, ambas prácticas presentan contextos de producción y propósitos comunes. Además, examinar rasgos relacionados con la simplificación en distintos conjuntos de textos traducidos al inglés desde distintos idiomas permite verificar ulteriormente la hipótesis de la simplificación como universal de traducción. Para el análisis, se compiló un archivo compuesto por cuatro corpus comparables de textos ingleses que reúnen versiones originales y traducidas de textos académicos y no académicos, respectivamente. Tras determinar índices de simplificación concretos (variedad léxica, densidad léxica, extensión promedio de las oraciones, presencia de oraciones subordinadas y de oraciones no finitas), se empleó una metodología de corpus para identificar cada índice en cada conjunto de textos en examen, de acuerdo con la hipótesis que ambas prácticas presentan los mismos índices de simplificación en la misma proporción. La comparación de los resultados refuta ambas hipótesis demostrando que cada práctica presenta índices de simplificación distintos y que, en todos los conjuntos analizados, los originales resultan ser más simples de las traducciones. Asimismo, la distribución de los rasgos examinados entre los distintos corpus parece seguir un patrón regular que representa un interesante punto de partida para estudios futuros.

Palabras clave: Estudios de traducción basados en corpus; Universales de traducción; Simplificación; Discurso académico; Contexto de producción de los textos.

Sumario: 1. Introduction. 2. Theoretical frame. 2.1. Academic Writing: an undefined textual practice. 2.2. Translationuniversals and the hypothesis of simplifications: an open debate. 3. Methodological frame. 3.1. Objectives and hypotheses. 3.2. Analyzed Corpora. 3.3. Methodology. 4. Results and discussion. 5. Final remarks.

1. Introduction

This paper aims to investigate the relationship between the context of production of texts and their specific textual features. To do that, it discusses some characteristics common to translation and Academic Writing (AW) as, despite their different natures, both practices are produced in a similar context involving common purposes and activities. First, the purpose of both is inherently explanatory: academic texts provide the readers with conceptual explanations about a specific topic, and translations offer a linguistic explanation of the source's content by making accessible a text written in an unfamiliar language. Second, both are metalinguistic and metatextual processes, with AW focusing on academic activities that rely on language to be accomplished, such as "educating students, [...]"

disseminating ideas and constructing knowledge” (Hyland 2009: 1), and translation using a target language to explain the original message. Third, both AW and translation are produced in specific social occasions in which they are created and used by the participants. Texts, including translations (Baker, 1996: 176), are influenced by the context in which they are produced (Palmquist, 1994-2012) and such influence is particularly evident in the case of the academic context (Motta-Roth, 2009). Thus, similar contexts of production make suppose the existence of similarities in the textual products resulting from both translation and AW.

Previous studies which analyzed separately translated texts and AW indicate that both practices involve a process of syntactic and stylistic simplification. Although AW is commonly thought of as elaborate and explicit, Biber and Gray (2010) demonstrate that its structures can be considered “simpler” than spoken conversations based on their syntactic structures. The results of their large-scale study comparing the structures of both types of discourse in English show that AW employs fewer subordinate clauses, particularly finite dependent ones, instead preferring condensed structures constituted by phrasal modifiers embedded in noun phrases. In this sense, Biber and Gray (2010) describe AW as “structurally ‘compressed’” rather than as “structurally ‘elaborated’” (Biber and Gray 2010: 2). Furthermore, according to the translation universals (TU) hypothesis of simplification (Baker 1993), translated texts are simpler than the originals from every perspective: lexically, syntactically and stylistically (Xiao and Yue 2009: 250). Concretely, from a syntactic point of view, translations exhibit more unconnected independent clauses than complex sentences, including secondary dependent clauses, that is, simpler clausal relationships.

Starting from the results about syntactic and stylistic simplicity, until now obtained only for each separate textual activity, this paper aims to provide a comparative description of the specific syntactic and stylistic features used to achieve simplification in translation and AW, according to the initial hypothesis that, as similar practices originated in similar contexts, both translation and AW will use similar textual features. Concretely, the analysis focuses on English translations and English academic texts, assuming the presence of concrete features common to any text translated into a certain target language, regardless its source language (according to the TU hypothesis explained in section 2.3), and the existence of generalizable characteristics in academic texts within the same linguistic community (see section 2.1 below).

The paper demonstrates several outcomes that will contribute to linguistics and translation studies in the following ways: exploring the

influence of the context of production on the textual products, providing a descriptive and comparative analysis of the specific features related to syntactic and stylistic simplification in translation and AW and further demonstrating the applicability of the TU hypothesis of simplification. In what follows, a theoretical framework is presented, resuming the multiple definitions of AW and pointing out the lack of specific and unanimous descriptions of its specific features. The paper follows describing the state of the ongoing debate about TU, their (in)existence and applicability, with a particular emphasis on the simplification hypothesis and its operationalization. Next, after a brief explanation of the advantages and disadvantages of corpus-based translation studies, some specific studies are discussed that relate to the topic of this paper, in order to frame it within the current state of the discipline and demonstrate the usefulness of a corpus-based methodology for the present research. The paper then describes its objective and hypothesis, explains the adopted methodology, and, finally, presents and discusses the results. The paper closes with some final remarks about the results and their implications in the discipline from a theoretical and methodological perspective.

2. Theoretical frame

2.1. *Academic Writing: an undefined textual practice*

No unequivocal and determinate definition of AW is unanimously accepted by the academic community (Flower 1990: 8). The definitions that do exist for AW are still very general; for instance, Hyland (2009: 1) considers AW “the way of thinking and using the language which exist in the academy”. Similarly, researchers have not reached a conclusive agreement about its distinctive features (Thonney 2011: 347). Scholars involved in the debate occupy three different positions. Some refute the possibility of identifying generalizable conventions among academic texts, as “writing” is not

an autonomous skill, generalizable to all activity systems [...] [it] does not exist apart from its uses, for it is a tool for accomplishing object(ive)s beyond itself. The tool is continually transformed by its use into myriad and always-changing genres [...] and [...] can be used for any number of object(ive)s [...] and transformed into any number of forms (genres). (Russell 1995, 58-59).

In contrast, other scholars argue in support of generalizable conventions that unite AW, regardless of discipline (Flower 1990, Thonney 2001) and maintain that these conventions can be taught as “a definite rule-set” (Lynch-

Biniek 2009). For instance, Thonney (2011) analyzed 24 academic articles in English and identified their common features. She (Thonney, 2011) identified several similarities among their structures, lexical choices, and tone. Structurally, academic articles respond to theories previously advanced about the topic, present a clear statement about the value of the text, and announce the paper's organization. These articles also show a preference for the use of academic and discipline-specific vocabulary, and their authors tend to adopt a voice of authority even though they are aware that some readers might disagree with their positions. Finally, other scholars fall somewhere between these two positions, accepting the hypothesis of common characteristics shared only by academic texts of a specific language (Mauranen and Ventola 1996; Bennet 2007) or within a specific discipline and discourse community (Wardle 2009: 784). This paper supports the third hypothesis, accepting the existence of generalizable characteristics in academic texts within the same Anglophone linguistic community.

2.2. Translation universals and the hypothesis of simplification: an open debate

The hypothesis of Translation Universals, firstly proposed by Mona Baker (1993), suggests the existence of specific features intrinsic and distinctive of any translated text, regardless the language pair. Baker's (1993) original proposal includes five TU: explicitation, simplification, normalization, "levelling out" and a tendency to avoid repetitions. Explicitation refers to the tendency to add explanations and use a more explicit style in the target texts (Xiao and Yue 2009: 249); simplification implies the use of a simpler language in translations in respect to the original texts (Zanettin 2013: 22) from a lexical, syntactic or stylistic perspective (Xiao and Yue 2009: 250); normalization indicates the preference to follow target language conventions instead of the original source language patterns (Zanettin 2013: 23); "levelling out" is the tendency of translations to occupy a central position, as for certain specific features, in the continuum extending between source and target language features, making translated texts "more like each other" in respect to the texts of a comparable corpus of originals in the same target language (Baker 1996: 184); finally, the tendency to avoid repetition indicates a translator's tendency to omit or substitute the repeated source text material.

In subsequent years, TU hypothesis came to occupy a central position within the discipline and a number of authors have participated to the debate either by trying to corroborate the TU hypothesis or by searching for further

universal features in texts of different language pairs (Pápai, 2004; Mauranen 2004; Xiao, 2010). As a result, two new universal features were detected in translated texts: interference, or the evidence of source language features in the target text, and the “hypothesis of unique items”, maintaining that translations present less “untranslatable”, source language-specific items than the texts drafted originally in the same target language (Zanettin 2013: 23).

The current research focuses on the hypothesis of simplification. This hypothesis contends that translations tend to be simpler than original texts at all linguistic levels. Considering and assessing simplification from different perspectives, previous authors have proposed specific indicators which allow for its operationalization. Even if many of such indicators can be related to more than one universal (Baker, 1996: 179), among the concrete manifestations of lexical simplification are usually included: lower lexical density and lexical variety in translation rather than in original texts (Laviosa, 1998a); more hyperonyms, usually introduced in case of lexical inequivalences; a general approximation to the source text concepts, by means of unprecise circumlocutions to substitute high-level words and expressions, more common synonyms and paraphrases to explain conceptual, specific or culture-specific terminology (Blum and Levenston, 1978: 399).

Among the syntactic features characterizing translated texts, previous authors highlight a preference for finite structures (versus non-finite ones), a minor use of suspending periods (Vanderauwera, 1985, as cited in Laviosa 1998b: 288) and a greater readability in terms of sentences and words length (Redelinguys and Kruger, 2015). Finally, from a stylistic perspective, the most frequent characteristic related to simplification seems to be the avoidance of repetitions (Vanderauwera 1985, as cited in Laviosa 1998b: 289). Due to the frequency with which such feature was identified in translated texts of any language pair (Toury 1991: 188), some authors describe it as an indicator of stylistic simplicity whereas others consider it a separate TU hypothesis (Laviosa 1998b: 289). Other manifestations of stylistic simplification are a minor sentence length (Laviosa, 2002: 62; Xiao and Yue, 2009: 250), obtained by dividing large sentences and expressions with stronger punctuation (Malmkjaer, 1997, Xiao and Yue 2009: 250) or by substituting them with shorter collocations (Vanderauwera 1985, as cited in Laviosa 1998b: 289), and “speakability”, or “the ease of reading aloud” (Puurtinen, 2003: 395).

In subsequent years, TU hypotheses have generated a great debate between translation studies scholars who support their existence (Laviosa 1998a; Zanettin 2013) and those who, on the contrary, refute it (Tymockzco

1998). Particularly, the contradictory evidences obtained by the analyses of different language pairs classifies simplification as more controversial than other TU hypotheses (Redelinghuys and Kruger 2015: 298), a central topic in current debates about their existence. Actually, whether Laviosa (1997, 1998a) and Xiao (2010) corroborated the hypothesis by analyzing a set of translations from English into Italian and into Chinese, respectively, Corpas Pastor et al. (2008) obtained the opposite results from the examination of a set of English-Spanish translations.

Detractors of TU consider them either an overgeneralization, due to the multiple contexts in which a translation can be used, or an undergeneralization, relating the universal features to the constraints that affect any type of communication, including translation (Redelinghuys and Kruger 2015: 296). Among them, House (2008: 11) considers TU language universals applicable to any linguistic context, thus also to translation, as it involves linguistic activity. Other scholars also participate in the dispute about the (in)existence of TU, arguing for their application to other types of texts. Heltai (2010, as cited in Lanstyák and Heltai 2012: 113) defines TU as “general tendencies of constrained language production” originated by the use of specific techniques employed to resolve inequivalences and transcultural problems, thus, applicable, but not limited, to translation. Shortly thereafter, Lanstyák and Heltai (2012) reconsider the concept of TU from a twofold perspective: from a general point of view, the authors relate TU to general communication, including any type of linguistic contact; focusing concretely on certain TU, such as simplification, unique items and interference, Lanstyák and Heltai (2012) describe them as “language contact universals” that characterize any cross-linguistic context (e.g., bilingualism, translation, etc.). From a similar perspective, Gaspari and Bernardini (2010) offer a comparative study of translated and non-native texts, suggesting the label “mediation universals” (Gaspari and Bernardini 2010: 3–4). These proposals further support this case study, suggesting the possibility to identify TU related features in a greater range of texts, even if they were originally conceived in relation to translated texts.

Actually, taking each concrete TU hypothesis one-by-one, some previous studies seem to demonstrate that certain universals can be identified in specific texts other than translations. This is the case of Mattioli (2018), who compares a corpus of translated novels with a corpus of travel novels identifying more evidences of explicitation in travel literature than in literary translations. Additionally, Swallow (2020) contrasts translated and edited texts in terms of their tendency toward normalization, discovering that this feature also affects edited language.

2.3. *Corpus-based methodology and previous studies*

Corpus-based studies were born in the 1950s. However, they were applied to Translation Studies only twenty years later fostered by the improvements of computer science (Mattioli 2018: 98) and the new descriptive perspective influenced by Even-Zohar's (1990) polysystem theory. Since then, corpus-based translation studies (CBTS) became particularly productive for the description and analysis of language (Sánchez Pérez 1995: 8–9). Even if the results of corpus-based studies seem to be deeply affected by the compilation criteria and some authors criticized them because of the impossibility to generalize the results due to the inclusion in the analyzed corpus of a single translation for each original texts (Malmkjaer, 1998: 534), CBTS still offer several advantages. Among them, Tognini-Bonelli (2001: 55) considers CBTS very useful to identify regular patterns of a certain language or linguistic variety; Gandin (2009: 147) underlines their adequacy to determine specific collocations, terminology, grammatical structures and stylistic patterns, suggesting their use to examine specialized languages or discourses; and Xiao and Yue (2009: 237) highlight the fruitful results of their application to study translation process, products and functions. From a methodological perspective, Rojo (2002: 5) denotes CBTS great exhaustiveness offered by the possibility to examine any of the occurrences included in a corpus and Gandin (2009: 137) points out the greater representativeness related to CBTS automatic nature which allows for analyzing big sets of authentic texts in a relative short time.

Thanks to the abovementioned characteristics, which are particularly fruitful combined with the new perspectives adopted by the discipline, corpus-based methods found a fertile ground in Translation Studies. Their use become particularly frequent in relation to TU thanks to their adequacy to compare originals and translations (Nilsson, 2004; Xiao, 2010) and to describe specific features of a set of texts (Laviosa, 1998a; Mauranen, 2004).

Corpus-based methods were also adopted in some previous studies focused on the comparison between translation and other textual practices. Among these, Mattioli (2018) compares translated and travel novels departing from their shared intercultural context and their common function of cultural mediation. Other authors, instead, preferred to use different methods to approach the same topic. Van Dyck (1990), for example, has edited a volume to compare translation to “deterritorialized” minority writing in modern Greek, considering both practices from the subaltern role of their authors. Similarly, Snodgrass (1993) compares translation and picture naming in order to investigate bilingualism issues, bearing in mind that both practices imply a vocal response, are influenced by the speed of such a

response, and entail a realization that is based on the use of semantic knowledge (Snodgrass 1993: 86). Gurdial Singh (2017: 28) balances the translation process with second language learning, concluding that both are “iterative, cumulative, dichotomous, integrative, interactive, forward and backward-looking mental operations.”

On the contrary, as for comparisons between translation and AW, there are only a few isolated cases. After a thorough search, only two studies were found: neither involves simplification of any type. First, Martínez et al. (2008) depart from the communicative aspects of the practices of translation and AW to realize empirical research of a sixth-grade classroom, studying the improvement of students’ writing abilities through translation strategies. Second, from a totally different perspective, Gil-Bardají (2009) compares translation and academic production in conceptual terms, focusing on Arabic to show how translation can help to produce a specific academic discourse.

The advantages presented by CBTS for the analysis of specific discourse added to the number of previous corpus-based studies examining TU and comparing translation with different textual practices demonstrate the suitability of a corpus-based method for the present case study. Considering the disadvantages of the corpus-based methodology, a special attention was given to the compilation criteria of the examined corpora, considering representativeness from a quantitative and qualitative perspective, and to the design of the methodology, taking into account the actual possibilities of the currently available corpus tools.

3. Methodological frame

3.1 Objectives and hypotheses

The main goal of this research is comparing translation and AW in terms of simplification in order to identify similar syntactic and stylistic features between both textual practices. The main hypothesis posits that translation and AW, as practices originated in similar productive contexts, present the same indicators of simplification to the same extent. Moreover, the study follows two further aims: observing the tendency to simplification in the translation corpus analyzed in the present study as a further validation of the applicability of simplification hypothesis of TU and exploring the relationship between the specific features of a text and the context in which it is produced. To pursue such general objectives, four specific goals are determined:

1. Compiling representative corpora representing translation and AW

2. Determining specific indicators to operationalize syntactic and stylistic simplification
3. Identifying the determined indicators in each corpus
4. Comparing the results obtained from among the corpora examined

3.2 Analyzed Corpora

The study analyses the two textual practices of translation and AW in English language. Actually, the TU hypothesis maintains that translations present common features regardless the language pair (Baker, 1993) however, academic texts seem to present generalizable characteristics only when produced within the same linguistic community (see section 2.1 above). Even if the main object of the study is the comparison between translation and AW, four corpora are compiled and examined, two for each considered practice, including original and translated versions of the selected types of texts, respectively. Such an archive of corpora permits to realize a further comparison between original and translated texts of each type, offering more relevant results and allowing for the verification of the simplification TU hypothesis pursued as an additional objective of the research.

As a consequence, the examined archive of corpora is compounded by four different sets of texts: Original Texts Corpus (OTC), including 33,129 types and 1,012,879 tokens; Translated Texts Corpus (TTC), including 32,935 types and 1,057,200 tokens; Original Academic Texts Corpus (OATC), including 35,477 types and 1,189,596 tokens; Translated Academic Texts Corpus (TATC), including 36,375 types and 1,294,770 tokens

Each corpus was compiled according to the following compilation criteria:

- quantitative representativeness: each corpus includes at least 1,000,000 tokens;
- qualitative representativeness: each corpus presents a textual, disciplinary, and linguistic variety including texts from different (academic or non-academic) fields and, in the case of translations, proceeding from an ample gamut of source languages;
- publication date: the selected texts were published between 2000 and 2019;
- inclusion of entire texts;
- balance: each corpus presents a balanced quantity of tokens and a symmetrical distribution of the texts across different subtypes and fields.

The resulting archive of corpus is presented in Table 1 which also indicate the distribution of the selected texts among the different fields, areas or subgenres.

Translation		AW	
Original texts	Translated texts	Original academic texts	Translated academic texts
OTC	TTC	OATC	TATC
30 NEWSPAPER ARTICLES <ul style="list-style-type: none"> • 6 science • 6 culture • 6 society • 6 world news • 6 economics 	30 NEWSPAPER ARTICLES <ul style="list-style-type: none"> • 5 science • 5 culture • 5 society • 5 politics • 5 world news • 5 economics Source languages: DE, ES, FR, SL	15 ACADEMIC PAPERS <ul style="list-style-type: none"> • 3 humanities • 3 applied sciences • 3 formal sciences • 3 natural sciences • 3 social sciences 	15 ACADEMIC PAPERS <ul style="list-style-type: none"> • 3 humanities • 3 applied sciences • 3 formal sciences • 3 natural sciences • 3 social sciences Source languages: DE, ES, FR, HE, IT, KOR, PT, RU, ZH
10 NOVELS	10 NOVELS	5 TEXTBOOKS <ul style="list-style-type: none"> • 1 humanities • 1 applied sciences • 1 formal sciences • 1 natural sciences • 1 social sciences 5 PhD DISSERTATIONS <ul style="list-style-type: none"> • 1 humanities • 1 applied sciences • 1 formal sciences • 1 natural sciences • 1 social sciences 	5 TEXTBOOKS <ul style="list-style-type: none"> • 1 humanities • 1 applied sciences • 1 formal sciences • 1 natural sciences • 1 social sciences Source languages: CZ, DE, FR, IT 5 TEXTBOOKS/MONOGRAPHIC VOLUMES <ul style="list-style-type: none"> • 1 humanities • 1 applied sciences • 1 formal sciences • 1 natural sciences • 1 social sciences Source languages: FR, DE, DK

Table 1. Distribution of the texts among the different corpora

During the compilation process, general texts were found to be shorter than academic ones. Thus, less (longer) texts were selected for the sets of texts representing AW in order to include at least 1,000,000 tokens in each corpus, accomplishing with the criteria of quantitative representativeness and balance. As a result, each corpus representing translation (OTC and TTC) includes 40 texts: 30 newspaper articles, distributed among five different fields (science, culture, society/politics, economics and world news) and selected from international newspapers of different countries, and ten novels of different subgenres (science fiction, historical, romance etc.) and written by renown authors from different countries. Each set of texts representing AW (OATC and TATC), instead, is compounded by 25 texts, equally distributed among five academic fields (humanities, applied sciences, formal sciences, natural sciences and social sciences): 15 scientific papers, 5 textbooks and 5 PhD dissertations or monographic volumes. Actually, the availability of PhD dissertations translated into English from a different source language is very limited, hence they were substituted by monographs or textbooks in the corpus of academic translated texts (TATC).

This paper studies translated texts from a target-oriented perspective, without considering the corresponding originals. Consequently, no parallel corpora are used and each textual practice is analyzed by examining only comparable corpora. For the same reason, the presence of different source languages in TTC and TATC (see Table 1) does not affect the comparability of the results. On the contrary, it permits an assessment of the validity of the TU hypothesis, according to which universal features are presented by any translation, regardless the language pair (Baker, 1993).

Once selected the texts to be compiled, indexes, abstracts in languages other than English, lists of references, and all images were eliminated and not considered during the analysis. The compilation of the archive of corpora presented in this section represents the successful pursue of the first specific objective of the research.

3.3 Methodology

The adopted corpus-based methodology consists of three main steps, corresponding to the three specific objectives left after the compilation of the archive of corpora, respectively:

1. Determining specific indicators to operationalize syntactic and stylistic simplification
2. Identifying the determined indicators in each corpus

3. Comparing the results obtained from among the corpora examined

In order to determine the specific indicators to operationalize syntactic and stylistic simplification, previous proposals were studied and combined. From the results of such inquiry, syntactic simplification seems to be related to the preference for simpler clausal relations, concretely realized by substituting non-finite structures with finite ones (Vanderauwera 1985, as cited in Laviosa 1998b: 288) and subordinate clauses with coordinative or independent unrelated ones (Biber and Gray 2010). As for stylistic simplification, it is assumed to be reached by opting for a minor fluency, lower lexical variety and lexical density, a preference for shorter sentences, with less repetitions and complex collocations (Laviosa 1998a). Resuming the previous postures, the specific indicators considered in the present study as concrete manifestations of syntactical and stylistic simplification are: (i) lexical variety; (ii) lexical density; (iii) mean sentence length; (iv) hypotactic structures; (v) non-finite clauses.

In the second step, each one of the determined indicators is identified in each corpus under analysis. To do that, different methods, programs and tools were used depending on the specific feature to be examined.

- i) Lexical variety: it was automatically calculated by WordSmith (Scott 2017) on the basis of the STTR (standardized types/token ratio) formula. STTR formula calculates the types/tokens ratio for each 1,000 words and averages the results for the entire text, that is, by standardizing the results for any 1,000 words (Redelinghuyis and Kruger 2015: 302).
- ii) Lexical density: it was calculated according to Stubbs's (1986) proposal. Once created a word list, lexical words were separated from the functional ones. The results of such separation were manually revised in order to spot and eliminate those ones which could represent a functional or a lexical word according to the context, for example, "once" which is a functional word in the role of conjunction, but a lexical word when used as adverb, noun or adjective. Only the items with a frequency equal to or greater than 50 were manually revised: such threshold permitted to avoid the results to be affected by high-frequency items while limiting the extensive, time-consuming manual revision. Then, the total number of lexical words was divided by the total number of tokens included in the corpus.
- iii) Mean sentence length: as the lexical variety, it was automatically calculated using a specific tool provided by WordSmith 7.0 (Scott 2017).

- iv) Presence of subordinate clauses: they were identified starting from the conjunctions or pronouns they are introduced by. In English, three types of subordinate clauses exist: those introduced by subordinative conjunctions, those introduced by relative pronouns and those which present no introductory item at all, which usually present a non-finite verb. As non-finite clauses were examined as a separate indicator (see point v below), here only the finite subordinative clauses were considered, that is, the ones introduced by subordinative conjunctions or relative pronouns. Consequently, once chosen an exhaustive list of conjunctions and pronouns (Several Authors, Mt San Jacinto College 2020), each one of them was searched for in the concordance list and its frequency was registered. Finally, the frequencies of each introductory item were added. As each of such item introduces a secondary clause, the total occurrences of introductory items correspond to the total number of secondary clauses.

However, not all the conjunctions and pronouns are actually used to introduce a secondary clause (e.g., “after” can be also used as a preposition, “what” can be used as a pronoun but also as an adjective, etc.). In order to distinguish the role of each examined pronoun or preposition and retrieve only those cases in which they are actually used as introductory items, further selection criteria were established, such as the presence of specific punctuation (e.g., *though* is used as a conjunction only between commas or immediately preceding the final full stop, otherwise it is usually a preposition), the position they occupy within the sentence (the pronoun “what” is used as an adjective only at the beginning of the clause, to signal a question), or the elements they are accompanied by (“because” is usually a conjunction however, when it is followed by the preposition “of” it accomplishes the role of a preposition, meaning “for the reason of”).

A specific methodology was used for the relative pronoun “that”. In fact, “that” is a polysemic word which can accomplish many different roles impossible to be distinguished manually because of the very high frequency of the word. According to the Grammar of English of the Cambridge Dictionary online (Several Authors, Cambridge University Press 2020), the pronoun “that” is used to introduce relative clauses in the four specific contexts presented in the following Table 2 (where “*” matches any character zero or more times and “\w+ {2,}” two or more words). In the analysis, specific regular expressions or wildcards were employed to identify such contexts automatically.

Context in which “that” introduces relative clauses	Example	Expression(s) used to detect each context
Superlative + “that”	The <i>best</i> game <u>that</u> I ever seen	the most/fewest/best/worst * that
“that” + preposition	The toy <u>that</u> you play <i>with</i>	That \w+ {2,} + preposition
“that” accomplish the role of subject or object of the sentence	A lesson <u>that</u> I learnt some years ago	a/an/the/some/any * that
“that” substituting the pronouns “who”, “whom” or “which” - in informal registers (Several authors, Cambridge University Press, 2020)	We met somebody last night <u>that</u> [instead of who] were at the Spanish course with you.	Not considered (The texts examined in this case study do not present informal register)

Table 2. Specific contexts in which “that” is used as a pronoun to introduce a relative clause and expression(s) used for their detection

- v) Non-finite clauses: according to the previous literature (Huddleston and Pullum, 2005), English language allows for four types of non-finite constructions:
- to-infinitival: preposition “to” + base form of the verb (to play);
 - bare infinitival (zero infinitive): base form of the verb (play);
 - gerund participial: -ing form of the verb (playing);
 - past-participial: participle tense of the verb (played).

The bare infinitival normally follows modal verbs (e.g., “study” in “I should study”) or the verbs “let,” “make” and “help,” (e.g., “go” in “let her go”) (Several Authors, Cambridge University Press 2020). Consequently, the resulting verbal expressions do not indicate non-finite clauses, hence they are not relevant for the study. As a result, only the other three types of non-finite structures are examined. Each one of them was combined to each item included in the same list of conjunctions used in (iv) and each combination was searched for in the concordance list, allowing for observing each entry in its context, hence for eliminating any possible error. By using the wildcard “*”, meaning “any character zero or more times”, only one search per each combination (conjunction + non-finite structure) was realized. For example, the expression “after *ing” allowed for retrieving all the occurrences of gerund participial forms of any verb introduced by the preposition “after”. A special attention was given to the past-participial irregular forms which change according to the paradigm. With the aim to include as many

participles as possible in the study, after observing the list of paradigms, the following regular patterns were identified and included in the search:

- participles ending in -en (e.g., beaten, chosen, driven, etc.)
- participles ending in -ught (e.g., brought, caught, fought, etc.)
- participles ending in -own (e.g., known, shown, thrown, etc.)
- participles ending in -ft (e.g., bereft, left, gift, etc.)
- participles ending in -nt (e.g., meant, learnt, spent, etc.)
- participles ending in -pt (e.g., crept, kept, slept, etc.)
- participles ending in -lt (e.g., built, dealt, felt, etc.)

Finally, in the last step, the results obtained from the four examined corpora were compared. Concretely, they were contrasted according to three combinations: Original academic texts vs. Translated texts (OATC vs. TTC), Original texts vs. Translated texts (OTC vs. TTC) and Original academic texts vs. Translated academic texts (OATC vs. TATC). To consider the actual difference among the features examined in each corpus, the statistical significance of each comparison is calculated with the log likelihood (LL) statistical test. The p value was set at 0.001, consequently the threshold of 6.63 was accepted to consider significant the results.

4. Results and discussion

The outcomes of the multiple comparisons between different corpora offer several interesting results. The most relevant findings are those recognized by the comparison between the corpora representing translated texts (TTC) and academic texts (OATC). These results offer data related to the main hypothesis of this study and are presented in Table 3 below: the first column indicates the examined indicators, the second and the fourth columns report the values corresponding to each indicator in each corpus, and the third column shows the LL value calculated for those results expressed in number of occurrences. The small arrow following or preceding the LL value (< or >) points at the corpus that presents an overuse of the examined indicator in respect to the other.

Simplification indicator	TRANSLATED TEXTS (TTC)	< LL >	ORIGINAL ACADEMIC TEXTS (OATC)
Lexical variety	45.46		37.76
Mean sentence length (in words)	17.47		21.68
Lexical density	0.52		0.61

Hypotactic structures	34,578	< 8,264.46	15,967
Non-finite clauses (total)	2,460	237.78 >	3,885
To infinitival	503	< 27.86	375
Gerund participial	1,747	56.48 >	2,359
Past participial	210	656.95 >	1,151
Non-finite subordinate clauses	866	417.53 >	2,043
Non-finite coordinate clauses	1,594	< 5.53	1,842

Table 3. Results of the comparison between texts representing translation and AW (TTC vs. OATC)

The results in Table 3 show that indicators with the highest values are distributed complementarily across the two corpora. AW reveals a low level of lexical variety as well as fewer hypotactic structures. Translations exhibit shorter sentences, a lower lexical density, and fewer non-finite clauses. These results refute the main hypothesis of the study: the complementary distribution of the simplification indicators across the two analyzed corpora denotes that, even if both textual practices display evidences of simplification, the texts representing translation and AW do not present the same indicators of simplification to the same extent.

Once evaluated the initial hypothesis, the results obtained from the comparisons between the corpora representing original and translated texts of each practice (OTC vs. TTC and OATC vs. TATC, presented in Table 4 and Table 5, respectively) allow to reach the additional objectives of the study. In this sense, the data obtained seem to refute the TU hypothesis.

In fact, the comparison of original and translated non-academic texts (OTC vs. TTC), presented in Table 4 below, does not show a considerable difference. In both corpora, lexical variety (44.32 in OTC and 45.46 in TTC), lexical density (0.49 and 0.52, respectively), and mean sentence length (18.20 and 17.47, respectively) present very similar values. Additionally, the greater use of non-finite structures in original texts than in translated ones (LL: 12.17) is actually due only to the difference between gerund participial structures (i.e., -ing form) (LL: 23.67) and coordinate clauses (LL: 16.28). As a result, the only analyzed feature that clearly differs between translated and original non-academic texts is the greater quantity of hypotactic structures in translations. Consequently, the translated texts analyzed in this study tend to be more complex than the original ones, refuting the TU hypothesis of simplification.

Simplicity indicator	ORIGINALS TEXTS (OTC)	< LL >	TRANSLATED TEXTS (TTC)
Lexical variety	44.32		45.46
Mean sentence length (in words)	18.20		17.47
Lexical density	0.49		0.52
Hypotactic structures	18,414	3,572.46 >	34,578
Non-finite clauses (total)	2,467	< 12.17	2,460
To infinitival	443	0.22 >	486
Gerund participial	1,865	< 23.67	1,568
Past participial	159	< 3.01	210
Non-finite subordinate clauses	800	< 0.12	866
Non-finite coordinate clauses	1,667	< 16.28	1,594

Table 4. Results of the comparison between original and translated non-academic texts (OTC vs. TTC)

In the corpora representing AW, instead, the difference between translated and original texts is greater than in the previous comparison. This appears in almost all the examined indicators, with the exception of lexical density, which does not show any substantial difference between the two corpora (0.56 and 0.61, respectively). Also in this case, the data refute the TU simplification hypothesis, showing that translations seem to be simpler than original texts only in their lesser quantity of non-finite clauses (LL: 12.62), whereas they are more complex in terms of longer sentences (26.29 vis-à-vis 21.68 words), lexical variety (38.95 instead of 37.76) and quantity of hypotactic structures (LL: 812.20). These outcomes are detailed in Table 5.

Simplification indicator	ORIGINAL ACADEMIC TEXTS (OATC)	< LL >	TRANSLATED ACADEMIC TEXTS (TATC)
Lexical variety	37.76		38.95
Mean sentence length (in words)	21.68		26.92
Lexical density	0.61		0.56
Hypotactic structures	15,967	812.20 >	23,244

Non-finite clauses (total)	3,885	< 12.62	3,870
To infinitival	375	18.84 >	545
Gerund participial	2,359	< 2.82	2,446
Past participial	1,151	< 63.24	879
Non-finite subordinate clauses	2,043	< 81.82	1,650
Non-finite coordinate clauses	1,842	10.49 >	2,220

Table 5. Results of the comparison between original and translated academic texts (OATC vs. TATC)

These results indicate that the TU hypothesis of simplification seems to be refuted in this particular case, or, at least, to be valid only for certain specific indicators. In fact, in each comparison, original texts present more features related to simplicity than translations. Secondly, both corpora representing translated texts (TTC and TATC) are characterized by different indicators of simplification, suggesting that simplicity varies according to the specificity of the discourse (academic or not) instead of the original or translated nature of the texts. Such conclusions are further underpinned by the degree of difference that the previous comparisons show. For example, the confrontations between original and translated non-academic texts (OTC vs. TTC) and AW (OATC vs. TATC), respectively, show very small differences between the values of the examined indicators (see Table 6), while the contrast between translation and AW (TTC vs. OATC) presents substantial differences with respect to almost all the considered simplification indicators.

Comparison Simplification indicator	TTC vs. OATC (Table 3)	OTC vs. TTC (Table 4)	OATC vs. TATC (Table 5)
Lexical variety	16.37	0.81	1.19
Lexical density	0.09	0.03	0.05
Mean sentence length (in words)	5.53	2.05	5.24
Hypotactic structures	LL: 1,124.59	LL: 106.97	LL: 903.20
Non-finite clauses	LL: 280.57	LL: 24.56	LL: 12.62

Table 6. Differences arising from each realized comparison

Finally, from the observation of all the comparisons discussed in the study, a regular pattern can be identified. The results shown in Table 7 highlight a very similar complementary distribution of the five considered indicators between the two corpora reviewed in each comparison.

Comparison	Translations vs. AW (Table 3)		Non-academic texts (Table 4)		Academic texts (Table 5)	
	TTC	OATC	OTC	TTC	OATC	TATC
Lexical variety	+	-	-	+	-	+
Hypotactic structures	+	-	-	+	-	+
Mean sentence length	-	+	+	-	-	+
Non-finite clauses	-	+	+	-	+	-
Lexical density	-	+	-	+	+	-

Table 7. Distribution of the examined features in each comparison

As the cases highlighted in grey in Table 7 demonstrate, in every comparison, a greater level of lexical variety and a greater quantity of hypotactic structures are opposed to a lesser use of non-finite clauses. This tendency suggests a further inquiry: Is the distribution of such features complementary? Concretely: Does a greater use of hypotactic structures imply a lesser use of non-finite clauses? Does a greater lexical variety imply the use of more hypotactic structures and less non-finite clauses? In every comparison performed in this study, these features are never identified in the same corpus at once. These interrogations and those resulting from any possible correlation among the distribution of the examined features could be used as points of departure or new hypotheses for further studies.

5. Final remarks

This paper presented a descriptive study comparing specific textual features in two different textual practices produced in a similar context, exploring the possibility that common contexts of production entail common textual features. To verify such idea, a specific case study was realized. The study aimed to compare translation and AW, as practices with similar contexts of production and purposes, in order to identify similarities as for the features related to simplification departed from the hypothesis that texts representing translation and AW exhibit the same simplification indicators to the same extent.

The study also followed two further goals: validating the TU hypotheses of simplification in the analyzed sets of texts and exploring the relationship between the specific features of a text and the context in which it is produced. In order to reach such objectives, a four-step methodology was designed and followed. Firstly, an archive of comparable corpora representing the practices of translation and AW were compiled; secondly,

five specific indicators were determined to operationalize syntactic and stylistic simplification starting from previous proposals (lexical variety, lexical density, mean sentence length, presence of hypotactic structures and of non-finite clauses); thirdly, each indicator of simplification was examined for separately in each corpus; and, finally, the results obtained were compared contrasting translation and AW, as well as the original and translated versions of academic and non-academic texts, respectively.

The outcomes refute the study's main hypothesis, showing that the texts representing the practices of translation and AW examined in this study do not exhibit the same indicators of simplification at the same extent, although their indicators follow a complementary distribution, instead (translations exhibit fewer non-finite clauses, a lower lexical density, and shorter sentences, while academic texts contain a lower degree of lexical variety and fewer hypotactic structures).

The comparisons between originals and translations of academic and non-academic texts allow to further verify the hypothesis of the TU of simplification that, in this study, has been refuted. Actually, tendency to simplification seems to depend more on the type of discourse (academic or non-academic) than on the original or translated nature of the texts. Even if the results of this exploratory study seem to discard the existence of a relationship between the context in which a text is produced and its specific features, further conclusions were drawn from the obtained outcomes. An overall observation of the results obtained from every comparison underlines a regular and complementary pattern in the distribution of the examined indicators. These outcomes suggest the potential existence of positive and negative correlations among the different features related to simplification, which offers an interesting departure point for future studies. Additionally, the primary results obtained in the present explorative study could be refined by realizing statistical examinations to take into consideration the great variability of the texts included in the four corpora, particularly evident in the first, main comparison (TTC vs. OATC, see Table 4) in which two different variables are crossed.

From a methodological perspective, the study contributes to the discipline by offering a replicable corpus-based methodology to assess concretely the simplicity or complexity of a corpus of texts. This method can be fruitfully used to analyze other sets of texts from different genres, contexts and types, or as an inspiration to design further similar corpus-based methodologies to identify semiautomatically different textual features, related or not to the TU hypothesis. In this sense, the current project encourages comparative CBTS from a twofold perspective. From the point of view of translation studies, it reviews traditional topics of the discipline proposing new methods to

investigate them. Concretely, this paper aims to participate to the still-open debates about the (in)existence of TU. From an interdisciplinary perspective, this study encourages further comparisons between translations and other types of texts, increasing the intrinsic interdisciplinarity of translation studies.

References

- ANTHONY, L. *AntConc* (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University, 2020. Available from <<http://www.laurenceanthony.net/software>>
- BAKER, M. (1993) "Corpus Linguistics and Translation Studies – Implications and Applications." In M. Baker, G. Francis & E. Tognini-Bonelli (Eds). *Text and Technology: In Honour of John Sinclair* (pp. 233-250). Amsterdam and Philadelphia: John Benjamins Publishing.
- BAKER, M. (1996) "Corpus-based Translation Studies: The Challenges that Lie ahead." In H.L. Somers (ed.) *Terminology, LSP, and Translation: Studies in Language Engineering in Honour of Juan C. Sager* [Benjamins Translation Library 18] (pp. 175-186). Amsterdam and Philadelphia: John Benjamins Publishing.
- BIBER, D. & B. GRAY (2010) "Challenging Stereotypes about Academic Writing: Complexity, Elaboration, Explicitness." *Journal of English for Academic Purposes*, 9:1, 2-20.
- BLUM, S. & E. A. LEVENSTON (1978) "Universals of Lexical Simplification." *Language learning*, 28:2, 399-415.
- CHESTERMAN, A. (2004) "Beyond the Particular." In A. Mauranen & P. Kujamäki (Eds). *Translation Universals: Do they Exist?* (33-49) Amsterdam and Philadelphia: John Benjamins Publishing.
- CORPAS PASTOR, G., R. MITKOV, N. AFZAL, & V. PEKAR. (2008) "Translation Universals: Do they Exist? A Corpus-based NLP Study of Convergence and Simplification." In *MT at Work. Proceedings of the Eighth Conference of the Association for Machine Translation in the Americas* (75-81). Waikiki, Hawai'i, USA 21-25 October 2008.
- EVEN-ZOHAR, I. (1990) "Introduction to Polysystem Studies." *Poetics Today*, 11:1, 1-6.
- FLOWER, L. (1990) "Negotiating Academic Discourse." In L. Flower, V. Stein, J. Ackerman, M. Kantz, K. McCormick & W. C. Peck. *Reading-to-write: Exploring a Cognitive and Social Process* (222-252). New York: Oxford University Press.
- GANDIN, S. (2009) "Linguistica dei Corpora e Traduzione: Definizioni, Criteri di Compilazione e Implicazioni di Ricerca dei Corpora Paralleli". *AnnaI/SS*, 5, 133-152.
- GASPARI, F. & S. BERNARDINI. "Comparing Non-native and Translated Language: Monolingual Comparable Corpora with a Twist". [Online

- version] 2010 [Last visit 18/10/2021]. Available from <https://www.academia.edu/download/46432454/Gaspari_and_Bernardini.pdf>
- GIL-BARDAJÍ, A. (2009) "Academic Discourse and Translation from Arabic: A case Study from the Spanish Tradition". *Babel*, 55:4, 381-393.
- GURDIAL SINGH, K. K. (2017) "Parallelism between Language Learning and Translating." *Journal of Modern Languages*, 17:1, 17-30.
- HOUSE, J. (2008) Beyond Intervention: Universals in Translation? *Transkom* 1:1, pp. 6-19
- HUDDLESTON, R. & G. K. PULLUM. (2005) *A Student's Introduction to English Grammar*. Cambridge: Cambridge University Press.
- HYLAND, K. (2009) *Academic Discourse: English in a Global Context*. London and New York: Continuum.
- KENNY, D. (2001) *Lexis and Creativity in Translation. A Corpus-based Approach*. Manchester: St. Jerome.
- LANSTYÁK, I. & P. HELTAI. (2012) "Universals in Language Contact and Translation." *Across Languages and Cultures*, 13:1, 99-121.
- LAVIOSA, S. (1997) "How Comparable Can Comparable Corpora Be?" *Target*, 9:2, 289-319.
- LAVIOSA, S. (1998a) "Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose." *META: Translators' Journal*, 43:4, 557-570.
- LAVIOSA, S. (1998b) "Universals of Translation." In M. Baker, K. Malmkjaer & G. Saldanha (Eds). *Routledge Encyclopedia of Translation Studies* (pp. 288-291). Abingdon and New York: Routledge.
- LAVIOSA, S. (2002). *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam and Atlanta: Rodopi.
- LYNCH-BINIEK, A. (2009) "Filling in the Blanks: They Say, I Say, and the Persistence of Formalism." [Online version]. *CEA Forum. College English Association*, 38:2, Summer-Fall, 2009. [Last visit 18/10/2021]. Available from <<http://www.cea-web.org>>.
- MALMKJAER, K. (1997) "Punctuation in Hans Christian Andersen's Stories and in their Translations into English." In F. Poyato (Ed.). *Nonverbal Communication and Translation: New perspectives and Challenges in Literature, Interpretation and the Media* (151-162). Amsterdam and Philadelphia: John Benjamins.
- MALMKJAER, K. (1998) "Love thy Neighbour: Will Parallel Corpora Endear Linguists to Translators?" *Meta: Translators' Journal*, 43:4, 534-541.
- MARTÍNEZ, R. A., M. FAULSTICH ORELLANA, M. PACHECO & P. CARBONE. (2008) "Found in Translation: Connecting Translating Experiences to Academic Writing." *Language Arts*, 85:6, 421-431.

- MATTIOLI, V. (2018) Los extranjerismos como referentes culturales en la literatura traducida y la literatura de viajes: propuesta metodológica y análisis traductológico basado en corpus. [Unpublished Ph.D. thesis]. Castellón de la Plana: Universitat Jaume I, 2018. [Last visit 18/10/2021]. Available from <<https://www.tdx.cat/handle/10803/587106>>.
- MAURANEN, A. (2004) "Corpora, Universals and Interference." In A. Mauranen & P. Kujamäki (Eds.). *Translation universals: Do they exist?* (65-82). Amsterdam and Philadelphia: John Benjamins Publishing.
- MAURANEN, A. & E. VENTOLA. (1996) *Academic Writing: Intercultural and Textual Issues* [Pragmatics and Beyond New Series]. Amsterdam and Philadelphia: John Benjamins Publishing.
- MOTTA-ROTH, D. (2009) "The Role of Context in Academic Text Production and Writing Pedagogy." In C. Bazerman, A. Bonini & Figueiredo (Eds.). *Genre in a Changing World* (317-336). Fort Collins: The WAC Clearinghouse Press.
- NILSSON, P. O. (2004) "Translation-specific Lexicogrammar? Characteristic Lexical and Collocational Patterning in Swedish Texts Translated from English." In A. Mauranen & P. Kujamäki (Eds.) *Translation universals: Do they Exist?* (129-142). Amsterdam and Philadelphia: John Benjamins Publishing.
- PALMQUIST, M. *Understanding Writing Situations*. [Online version]. Writing@CSU. Colorado State University. 1994 - 2012 [Last visit 18/10/2021] Available from <<https://writing.colostate.edu/guides/guide.cfm?guideid=3>>
- PÁPAI, V. (2004). "Explicitation: A Universal of Translated Text?" In A. Mauranen & P. Kujamäki (Eds.) *Translation Universals: Do they Exist?* (143-164). Amsterdam and Philadelphia: John Benjamins Publishing.
- PUURTINEN, T. (2003) "Genre-specific Features of Translationese? Linguistic Differences between Translated and Non-translated Finnish Childrens Literature." *Literary and Linguistic Computing*, 18:4, 389-406.
- REDELINGHUYS, K. & H. KRUGER. (2015) "Using the Features of Translated Language to Investigate Translation Expertise: A Corpus-based Study." *International Journal of Corpus Linguistics*, 20:3, 293-325.
- ROJO, G. "Sobre la Lingüística Basada en el Análisis de Corpus." [Online version]. *UZEI Hizkuntza-corporak. Oraina eta geroa*, 2002-10-24/25. [Last visit 18/10/2021]. Available from <<http://uzei.com/Modulos/UsuariosFtp/Conexion/archivos54A.pdf>>

- RUSSELL, D. R. (1995) "Activity Theory and Its Implications for Writing Instruction." In D. Russell & J. Petraglia (Eds.). *Reconceiving Writing, Rethinking Writing Instruction* (51-77). Mahwah: Erlbaum.
- SÁNCHEZ PÉREZ, A. (1995) "Definición e Historia de los Corpus." In Sánchez Pérez, A. *CUMBRE: Corpus Lingüístico del Español - Contemporáneo* (7-24). Madrid: SGEL.
- SCOTT, M. (2017) *WordSmith Tools* version 7, [computer software] Stroud: Lexical Analysis Software. Available from <<http://lexically.net/wordsmith/>>
- Several Authors (Cambridge University Press). *Cambridge Grammar of Cambridge Dictionary*. [Online version]. 2020. [Last visit 18/10/2021]. Available from <<https://dictionary.cambridge.org/grammar/british-grammar/>>
- Several Authors (St Mt San Jacinto College). "Prepositions and Conjunctions." [Online version]. 2020. [Last visit 18/10/2021]. Available from <<https://www.msjc.edu/learningresourcecenter/mvc/mvcwritingcenter/documents/handouts/Prepositions-and-Conjunctions.pdf>>
- SNODGRASS, J. G. (1993) "Translating versus Picture Naming Differences and Similarities." In Schreuder, R. & B. Weltens. *The Bilingual Lexicon* (83-114). Amsterdam and Philadelphia: John Benjamins Publishing.
- STUBBS, M. (1986) "Lexical Density: A Technique and Some Findings." *Talking about text*, 27-48.
- SWALLOW, H. (2020) "Edited vs Translated English: Normalisation in Mediated Language through the Lens of the Split Infinitive." In S. Granger, M. A. Lefer (Eds.) *Translating and Comparing Languages: Corpus-based Insights*. Corpora and Language in Use Proceedings 6 (191-211). Luvain-la-Neuve: Presses universitaires de Louvain.
- THONNEY, T. (2011) "Teaching the Conventions of Academic Discourse." *Teaching English in the Two-Year College*, 38:4, 347-362.
- TOGNINI-BONELLI, E. (2001) *Corpus Linguistics at Work*, Amsterdam and Philadelphia: John Benjamins Publishing.
- TOURY, G. (1991) "What are Descriptive Studies in Translation Likely to Yield Apart from Isolated Description?" In K. Leuven-Zwart, T. Naaijken & A. B. M. Naaijken (Eds). *Translation Studies: The State of the Art. Proceedings from the First James S. Holmes Symposium on Translation Studies* (179-192). Amsterdam: Rodopi.
- TYMOCKZCO, M. (1998) "Computerized Corpora and the Future of Translation Studies." *Meta: Translators' Journal*, 43:4, 652-660.
- VAN DYCK, K. (1990) "Introduction to Translation and Deterritorialization." *Journal of Modern Greek Studies*, 8:2, 169-171.

- VENUTI, L. (1995) *The Translator's Invisibility*. London and New York: Routledge.
- WARDLE, E. (2009) "'Mutt Genres' and the Goal of FYC: Can We Help Students Write the Genres of the University?" *College Composition and Communication*, 60:4, 765–89.
- XIAO, R. (2010) "How Different is Translated Chinese from Native Chinese? A Corpus-based Study of Translation Universals." *International Journal of Corpus Linguistics*, 15:1, 5-35.
- XIAO, R. & M. YUE. (2009) "Using Corpora in Translation Studies: The State of the Art." In P. Baker (Ed.) *Contemporary Corpus Linguistics* (237-262). London: Continuum.
- ZANETTIN, F. (2013) "Corpus Methods for Descriptive Translation Studies." *Procedia - Social and Behavioral Sciences*, 95, 20-32.