Genre-based Approach to Corpus Compilation for Translation Research

Jana Kegalj & Mirjana Borucinsky

University of Rijeka, Rijeka, CROATIA
Faculty of Maritime Studies

Abstract

Translation research focuses mainly on parallel and comparable corpora, whereby it is constantly faced with issues of representativeness, balance and comparability as its main constraints. This research aims to introduce the concept of genre as a way of observing linguistic features under controlled conditions. The study analyses the application of external and internal criteria with particular focus on the genre criterion in selecting texts for the compilation of a highly-specialized bilingual maritime legal corpus, consisting of source texts in English and their translations into Croatian. The main advantages and constraints of genre as a criterion are discussed. The main benefits of such an approach are found in its application in translator training and practice. In addition, genre-based approaches to corpus analysis may raise awareness of generic features specific to a target language, ultimately improving the quality of translation.

Keywords: translation studies, genre, corpus research.

1. Introduction

Nowadays, corpus tools are extensively used to study specific features of translated texts. This significant advance in translation studies brought about by corpus linguistics has provided an empirical basis for translation research, opening up new perspectives for the scientific study of translation. In addition to this empirical turn (Snell-Hornby, 2010), translation studies as a discipline have also undergone a cultural turn (*Ibid.*) whereby translations have started to be viewed as a linguistic phenomenon and as a cultural and social fact. This has had implications both for translator education, implying that translators have had to develop various competencies besides the linguistic competence, namely communicative, cultural, translation and generic competencies (see Antunović, 2003), and translation research, placing a text within a context and setting it against other similar texts, thus providing a broader scope for investigation. One of the significant advantages of understanding translation as an act of communication and a cultural phenomenon is the consideration of the function and purpose that the text has in the target language (TL), which has to correspond to the function and purpose of the source text (ST) in the source language (SL).

- Corpus tools are extensively used in translation studies.

- Genre represents a useful selection criterion in corpus compilation.

- Genre-based selection of texts may positively influence issues such as corpus representativeness, balance and comparability.

- Genre-based approach to translation research may benefit both translator training and the scientific study of translation.

This is related to the concept of genre, which is also viewed as a communicative event (Swales, 1986) and a conventional form of text (Hatim & Mason, 1990), indicating a cultural and communicative purpose and having certain specific rhetorical purpose features. In their translation model, Hatim and Mason (1990) specifically identified genre as the primary consideration in the process of translation influencing the choices that translators make. The genre factor also has implications for corpus study as corpora are sensitive to different text types. Therefore, an informed genre-based selection of texts may positively influence corpus representativeness, balance, and comparability issues. Furthermore, genre analysis is more than mere linguistic analysis, i.e., it deals with the way a text is composed, the reasons why some linguistic units were selected over others and the combinations of specific linguistic units in a particular context. It also contributes to understanding the form–function correlation and the cognitive structuring of information (see Bhatia, 1993).

Following the abovementioned considerations, a specialized corpus was compiled for this research, emphasizing genre during text selection and considering how genre influences its representativeness and balance of the corpus. The hypothesis is that by introducing the criterion of genre as a key factor, a corpus is placed in a controlled environment, thus getting relatively closer to the ideal of corpus representativeness.

2. Theoretical considerations

As defined by Swales (1981), genre is "*a recognizable communicative event characterized by a set of communicative purpose(s) identified and mutually understood by the members of the professional or academic community in which it regularly occurs*." Therefore, genre analysis represents a holistic approach to discourse, investigating its distinctive features and emphasizing its communicative function. In translation studies, genre has introduced the idea of observing text function and purpose in the TL, which does not necessarily have to coincide with the function and purpose of the source text in the SL. Therefore, in translation, a translator departs from a text in a particular genre, translating it into the TL genre complying with the norms of the target culture. In this process, a translator needs to have generic competence to translate a particular text into the TL genre. Thus, genre analysis in translation studies connects the microlevel of the text with the macrolevel of discourse and context, integrating the cognitive, social and professional approach to translation (García Izquierdo & Montalt i Resurrecció, 2002). García Izquierdo and Montalt Resurrecció (*Ibid.*) believe that a translator translates one genre into another, so the genre is the starting point and translation goal. In other words, it is the "interface" between language and its use in a particular situation.

Genre theory emerged and developed in linguistics at the same time when translation studies were established as an independent discipline, so as the interest for genre grew within discourse analysis, it was followed by a similar interest within translation studies (Biel, 2018). The first to recognize the importance of genre for translation studies was James (1989), who advocated its application in translator training. He also believed that translations were translations of other genres and that inexperienced translators must be systematically exposed to different genres. Shäffner (2000) claimed that genre analysis would help raise awareness in future translators of

language patterns on various levels, whereby they could understand the complexity of the translation process and develop their translation competence. Various research has also shown that translations performed by inexperienced translators are more stylistically marked, which results from their lack of knowledge on linguistic units used in a particular genre in a foreign language. Snell-Hornby (1988) also applied text typology in the translation process and emphasized that the more specialized a text is, the easier it is to define the purpose of translation, thus making the translation more adapted to the TL. Within his *skopos theory*, Vermeer (1989) also pointed out the function of generic norms that affect the translation process's decision-making. Reiss and Vermeer (2013) believed that selecting a translation strategy depended on translation function, as translators attempted to adjust the texts to the target culture's expectations. In that sense, Vermeer (*Ibid.*) concluded that genre could not be neglected as generic conventions influence the selection of language units, translation strategies, and the understanding of the text. Therefore, translation implies adjusting the target text to generic conventions of the target culture to the target audience's expectations. Bhatia (1993) described a multi-perspective model of genre analysis, which consists of seven steps, including the analysis of the context, i.e., the domain, discourse community, sender and receiver of the text, communicative purpose, medium, content, and extra-textual reality.

Genre is often associated with the concept of register and the differences between the two notions are not always clear-cut. For example, Biber and Conrad (2009) stated that register and genre differ because the register pays attention to content, not form like genre, and frequent lexicogrammatical features with a specific communicative function. According to the authors, the register is a more general, abstract category, while the genre is concrete. They also emphasized that genre analysis requires whole texts as all features are essential, including those that appear only in one part of the text but may be vital as they are conventionalized, while register analysis can be performed on excerpts. However, Biber himself used the term genre differently throughout his work and later opted for the term register as central to his theory.

Lee (2001), on the other hand, considers the concepts of register and genre as two distinct but complementary approaches to the same area, whereby register observes a text as a conventionalized, functional structure dependent on the social situation, while genre observes a text as a culturally recognizable structure or a group of texts with the same conventionalized features. According to the author, genres are realizations of registers, and each genre has certain lexicogrammatical and discourse features of one or more registers. Generally, Lee (*Ibid.*) believed that genre as a recognizable category might be a useful starting point in research. In this study, genre and register are considered distinctive but complementary notions, providing valuable insights into text features.

3. Method

As research has shown, genre plays an important role in translation. Hence, this paper introduces genre analysis as a significant methodological tool in translation studies and considers genre as a criterion in the corpus compilation process for purposes of translation research. Consequently, to investigate the role and effect of genre, a corpus was compiled based on genre as the main criterion. The selected genre, in this case, was defined very specifically as belonging to institutional, legal maritime English. The direction of translation was from SL English into TL Croatian. Firstly, considerations were made about the type of corpus suitable for translation research, and, secondly, the stages of corpus compilation were defined and discussed. Finally, significant issues arising during the compilation process were analyzed, and the advantages and disadvantages of such an approach were summarized.

3.1 *Corpora in TS*

Corpus linguistics aims to identify, describe and explain the patterns of language use, focusing on those frequent, typical, central, and expected ones and showing that language is structured in a particular way and that patterns are not random and accidental (see Biel, 2018; Biber & Jones, 2009; Stubbs, 2004), but relatively organized and structured. Although historically, translations were considered a deviation from the norm, i.e., a hybrid form of language influenced by a foreign language, they were not considered representative and were not included in the first corpora. However, Baker (1993) recognized the potential that corpus tools might have in developing the discipline of translation studies and advocated for the use of corpora in finding typical patterns in translations that might have universal features – the so-called translation universals. This provided translation research with an empirical basis and proved as a fruitful source of data about various aspects of translation, opening new research questions (Laviosa, 2004). As a result, corpus-based translation research has dealt with a range of subjects, some of the most prominent ones being translation norms, translation universals, and translation styles (Zanettin, 2000). Generally, corpora reduce subjectivity by providing authentic data extracted from a large source of linguistic material. However, the issues of representativeness and balance still need to be considered, and any observations based on corpus analyses are necessarily restricted to and by the sample under study.

Two types of corpora are primarily used in translation studies: parallel bilingual or multilingual corpora and comparable corpora, which are frequently combined in research (Biel, 2018). Parallel corpora consist of texts in the SL aligned with their translations in the TL. They may consist of texts in both directions of translation. As text alignment is rather time-consuming, these are not compiled often, although they provide a valuable insight into translation strategies, translation universals, or terminological issues. Biel (2018) distinguishes two types of comparable corpora: comparable monolingual corpora, containing texts originally produced in one language and translated texts, or bilingual or multilingual corpora, consisting of texts in different languages selected according to predefined criteria. The latter are also called translation-driven corpora (Zanettin, 2000) since they are made for translation research or education. Some examples of large corpora compiled for translation studies include the *Translational English Corpus* (TEC)[1], *European Parliament Proceedings Parallel Corpus*[2] (Europarl), *GENTT Corpus of Textual Genres for Translation*[3]. In order to improve the quality, validity, and reliability of the obtained results, this study combined corpus methods, i.e., applied corpus triangulation[4] (Malamatidou, 2018).

The corpus was designed to account for different needs and answer research questions, so the issue of its size and composition were also taken into consideration. The consensus is that a corpus should be as large as possible, and nowadays, electronic corpora enable the creation of such large quantities of linguistic data. However, in this case, several restrictions inevitably limited the size of our corpus, the first one being the defined type of corpus and the second the key criterion of genre, so a compromise had to be made regarding corpus size. As the corpus had to be a parallel one, texts translated into TL, in this case, Croatian, had to be found that satisfied the genre criterion. The availability of such texts emerged as one of the major limiting factors as not all texts are translated, and some are unavailable for confidentiality reasons (e.g., private legal texts, lawsuits). However, considering that specialized corpora are regarded as a rich source of

---

[1] Available at http://www.llc.manchester.ac.uk/ctis/research/english-corpus/.

[2] Available at https://www.statmt.org/europarl/.

[3] Available at http://www.gentt.uji.es/?q=es%2Fel-corpus.

[4] Triangulation implies combining methods to verify the obtained results through, for example, quantitative and qualitative data processing or combining different types of corpora like parallel and comparable corpora.

data, provided they are balanced and carefully compiled (Bowker & Pearson, 2002), size was not a significant issue in this case, so more attention was paid to its internal structure.

The internal structure of the corpus was inevitably directed towards something Biel (2010) calls legicentrism, i.e., it had to include only the genre of legislative texts, as other genres are not available owing to the reasons mentioned above. This is not exclusively the problem of this corpus, but all legal corpora. In general, legal corpora, such as this one, do not need to be very large owing to the conservativism and formulaicity of the legal language, so after a specific size of the corpus is reached, adding new texts to it does not contribute significantly to its representativeness or the results it yields (Bhatia et al., 2004). Some examples of legal translation corpora are JuriGenT⁵ (Vanden Bulcke) and JudGENTT⁶ (Borja Albi), which employ parallel and comparable corpora of legal texts for terminological purposes or generic research of translated texts. This approach is also advocated by Bhatia et al. (2004), who believe that such an approach results in a more relevant resource for legal language and legal translation research.

3.2 *Criteria for text selection*

Carefully defined criteria for selecting texts that are to be included in the corpus are important to ultimately design a corpus that is representative and balanced as much as possible, and that satisfies the needs of the study. In this case, the corpus compilation evolved through several stages:

1) **Planning**: in this stage, the type of corpus is defined, relevant criteria selected, and decisions made based on the criteria. The main criterion, in this case, was genre, but other criteria such as corpus size, type of corpus, and availability of texts were also taken into consideration.

2) **Primary filtering**: texts were collected following the criteria set in the previous stage.

3) **Secondary filtering**: the texts selected in the previous stage were further checked, and certain parts of the texts had to be omitted. For example, legal texts originally written in Croatian contained legal definitions of terms used in the act. These acts frequently refer to a directive or another legal instrument by an international body and take definitions from those acts. Therefore, the definitions represent translations of the said terms, and as such, would not be fit as a part of the comparable corpus as they represent translations.

4) **Text processing**: To conduct a corpus analysis, the texts needed to be prepared for computer processing. First, they had to be converted into a suitable format. In this case, the corpus was compiled in Sketch Engine (Kilgarriff, Rychlý, Smrz & Tugwell, 2004), which supports most file formats, but not all. Parallel corpora had to be aligned. Secondly, it was important to decide whether the alignment would be at sentence or paragraph level, which depends on the aim of the study, the nature of the texts, and the program used for corpus search. As the texts selected for this particular corpus are highly structured legal texts in which intertextuality, discourse features, and deictic markers play a significant role, it was decided to align the texts at the paragraph level to investigate such features. Another reason for such alignment was that SL and TL sentences do not have a 1:1 correspondence. The

---

⁵ JuriGenT is a bilingual corpus of legal texts in Dutch and Spanish, created by Patricia Vanden Bukcke and Carine De Groote in 2017 for terminographic research.

⁶ JudGENTT corpus of legal texts was developed by the group of scientists GENTT (http://www.gentt.uji.es/) at the University Jaume I in Spain.

_____

alignment was performed first using Memsource.[7] Memsource is a tool that also enables the alignment of documents. However, the alignment had to be checked manually since the program could not process larger texts with the precision necessary for corpus compilation.

The criteria for corpus compilation may vary according to the purpose of the corpus and research aims and are frequently a result of individual assessment of the compiler. The ultimate goal is to obtain a corpus as representative as possible for the language variety being studied. For example, according to the recommendations of the EAGLES[8], the criteria were divided into two groups: external non-linguistic criteria and internal linguistic criteria. External criteria would be defined socially, extralinguistically, regardless of the language structures involved, including literary genre, medium, style, and mode, while internal criteria include topic and style. The authors emphasized that these criteria are interconnected and concluded that both types of criteria should be applied equally in the selection. They recommended starting the selection process based on external criteria and then fine-filtering it through internal criteria. Then the process should be cyclically repeated until a stable sample is achieved (see Biber, 1993). Relying only on external criteria might neglect some significant variations among texts while using exclusively internal criteria might not offer data on the relation between a text and its context. The authors of the EAGLES project listed many possible criteria for text selection, but not all criteria can be used simultaneously in research. The researcher must choose those criteria relevant for their study and apply them consistently in corpus compilation to achieve representativeness, balance, and comparability as much as possible.

Based on the EAGLES description, it might be assumed that genre is a category that may be intuitively recognized and differentiated according to external criteria. Usually, this refers to information such as origin, author, intended audience, intended goals, socio-cultural context, historical background, and topic, which also corresponds to the steps in Bhatia's model of genre analysis (1993). This had implications for the corpus in this study as the genre could be defined much more specifically, while the criterion of topic was introduced as another important indicator. According to EAGLES, the criterion of a topic is somewhat elusive and open-ended but frequently used as it is easily recognized. For example, the genre of newspaper articles may have various topics. In our view, it might be used as an assisting criterion but not a decisive one.

4. Discussion

Genre has recently become increasingly recognized as an important criterion in text selection. On the one hand, this is because it can easily be identified according to external features, which are also relatively intuitive. These are culture-bound characteristics present in the minds of all the speakers of a specific language and so expected in a particular context. On the other hand, genre is also distinguished by internal features; specific linguistic units frequently found in a specific genre that differentiate it from other genres. These are all the reasons for introducing this concept in the corpus compilation process to make a corpus as representative and comparable as possible. In such a corpus, genres may be compared within one language, or the same genre may be observed across two languages. This has multiple implications for translation studies. It can be used in translator training to raise awareness in translator trainees about the nature of genres, their external characteristics, and frequent and expected linguistic features. It can also be used to

_____

[7] https://www.memsource.com/ – Memsource is a system for translation management that enables translators to create their translation memories and translation corpora to improve the speed and quality of translations.

[8] Project (http://www.ilc.cnr.it/EAGLES/home.html).

study translation strategies used to convey a text in a particular genre, the development of genres under the influence of translation from another language, or universal genre features.

Genre is also a category that includes other categories specified by the EAGLES group. For example, the text's origin, the sender and receiver of the message, the purpose and context. Therefore, filtering texts through the genre criterion automatically implies the consideration of these external criteria. The category of the topic also remains, which is complementary to the category of genre and may be considered alongside the genre. Depending on the goals of the research, a topic may be disregarded (e.g., studying the language of newspaper articles in general), it may be considered as another key criterion (e.g., studying the language of newspaper articles dealing with immigrants and refugees), or it may be an additional criterion to genre to create a more controlled environment for observing a specific language.

4.1 *Case study*

The corpus compiled for this study, the institutional, legal maritime English and Croatian corpus (MarLaw), was guided by the principles of genre theory. Therefore, besides the main criterion of genre, the topic criterion was introduced as an ancillary element. The topic criterion was introduced here despite its unpredictable nature. In this case, the topic restricted and determined the area of specialized vocabulary, which may serve to understand and compare terminology used in both languages and detect possible differences between translations and texts originally written in Croatian. This also enabled easier isolation of typical legal terms and phrases. However, after filtering the text according to the main genre criterion, the topic was not an essential criterion but a supplementary one.

The texts selected for the corpus were institutional, legal texts, i.e., conventions, acts, and regulations that govern maritime affairs on the national and international level. This genre is conventionalized and recognizable in its formal structure and composition with the usual preamble, the summary of objectives, the body consisting of articles, and the closing protocol.

The introduction of genre as the main guiding principle also brings about several issues. In the case of the corpus compiled for this study, the texts belonging to the genre can easily be identified and available online on the institutions' web pages that pass such legal instruments (i.e., International Maritime Organization). Furthermore, the countries which are members of these institutions organize the translation of the legal instruments into national languages, so their translations are also available on official web pages. Therefore, the compilation of the parallel corpus would include the selection of SL and TL texts and their alignment, which is a time-consuming process but yields a corpus that will provide insightful data about the translation process, translation strategies, terminology, and syntactical structuring. However, the compilation of the comparable corpus required further considerations. In order to be comparable, it was necessary to find texts with similar legal strength. The legal instruments closest to international conventions are national acts, while regulations have equivalent variants on the national level. Therefore, to be balanced and comparable, both legal instruments of similar legal strength should be represented with an approximately equal share in the parallel and comparable corpora.

However, in attempting to achieve comparability as much as possible, some authors (see Malamatidou, 2018; Leech, 2007) warn about reducing representativeness. A way to achieve the best of both categories is to set controlled conditions, such as genre, and apply the same criteria to both corpora. As genre analysis studies whole texts, this was the criterion used in corpus compilation, along with the criterion of size, i.e., making the corpora of approximately equal size.

Another issue was the issue of intertextuality. Specifically, national laws refer to international legal instruments, adopting the definitions of terms from them. Therefore, the comparable corpus could not contain these parts as they represent translations of international

documents, so they had to be omitted from the corpus. As for other parts of the acts, in consultation with a legal professional, laws are drafted according to national rules and customs, as well as linguistic standards, so even though they contain provisions by an international legal instrument (e.g., EU directive or similar), they are composed independently of them, i.e., they do not represent translations of these documents.

These considerations resulted in genre-driven parallel and comparable corpora representative of the institutional maritime legal genre (i.e., in two languages – English and Croatian), both containing approximately 500,000 tokens, which is considered appropriate for a specialized corpus. In addition, the compiled corpora can be used to conduct a further linguistic study regarding, for example, the nature of translations, the issue of translation universals, the possible differences of translated texts in relation to original texts, a comparison of legal language or genre across two different languages.

4.2 *Genre-based corpus approach*

Applying the genre criterion during the compilation of the MarLaw corpus resulted in creating a relatively balanced and comparable corpus representative of the legal maritime genre. Therefore, it can be used for a detailed description of the genre, the lexicogrammatical patterns used in the genre in two languages, compare the genre in two languages, compare the terminology used, study the nature of translated language (Croatian translations), observe the possible instances of translation universals, or detect translation strategies in this particular language pair within controlled conditions of genre and topic. Further, it can be used to compare this genre to other genres (e.g., EU legal language, the so-called Eurospeak or Eurojargon) and for educational purposes in translator training to raise awareness of specific genre features.

The genre-based approach has certain issues which have to be overcome on a case-to-case basis. For example, some genres may overlap, or their boundaries may not be as clear cut, affecting the corpus' representativeness. Furthermore, there may be significant differences in the case of comparison between languages, e.g., academic texts in German and Czech (see Čmejrkova, 1996), which makes it difficult to compile a parallel corpus. In the new environment of the World Wide Web, genres have undergone specific changes that have somewhat blurred the genre boundaries or introduced new genres. This criterion also significantly influences the size of the corpus, but this could be overcome by creating several subcorpora belonging to the same corpus, e.g., a corpus of legal texts consisting of a subcorpus of acts, a subcorpus of legal proceedings, a subcorpus of judgments and a subcorpus of pleas. Another disadvantage may be that such a process is more time-consuming than a simple collection of texts from the web, for example.

Regardless of the difficulties, a carefully conducted genre analysis undertaken prior to corpus compilation may significantly contribute to compiling a more representative and comparable corpus which could prove an exhaustive source for linguistic and contrastive research.

5. Conclusion

The paper tried to describe the advantages, disadvantages, and issues regarding the use of genre as a criterion for text selection in corpus compilation. It has been shown that corpora are sensitive to the category of text type and cannot be considered representative unless this feature is taken into consideration. Of course, representativeness is then restricted to that particular genre, but insights gained from such a corpus provide relevant results which may be compared with other genre-based corpora in the same or another language.

While in the SL, the genre is a result of prior consideration of the category that the text will belong to, the function it will have, and the situation in which it will be used, in translation,

_____

the translator starts from the genre, and then decides upon its context and function which ultimately affects the choice of linguistic elements used. In that case, the context of the situation and the function of a text in SL may not correspond to the context and function in the TL. This places genre as a central notion on the continuum from the source text to target text.

The paper has provided an overview of issues that arise in corpus compilation and which have to be considered in order to compile as representative, balanced and comparable corpus as possible. A case study has also shown how some of these issues may be overcome. Therefore, creating genre-based corpora has implications for translation training and translation studies, providing valuable insights into the relation between the text, its function, and its context and contributing to the understanding of cognitive structuring of information.

Acknowledgements

The authors declare no competing interests.

References

Antunović, G. (2003). *Remarks on (purported) translators' tasks and translation teaching*. *SRAZ* XLVII-XLVIII, 13-22.

Baker, M. (1993). Corpus linguistics and translation studies – Implications and applications. In M. Baker, F. Gilli & E. Tognini-Bonelli (Eds.), *Text and technology. In honor of John Sinclair* (pp. 233-250). Amsterdam: John Benjamins.

Bhatia, V. K. (1993). *Analysing genre - Language use in professional settings*. London: Routledge.

Bhatia, V. K., Vijay K., Langton, N. M. & Lung, J. (2004). Legal discourse: Opportunities and threats for corpus linguistics. In U. Connor & Th. A. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 203-231). Amsterdam/Philadelphia: John Benjamins.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, *8*(4), 243-257.

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge and New York: Cambridge University Press.

Biber, D., & Jones, J. K. (2009). Quantitative methods in corpus linguistics. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics. An international handbook*, Vol. 2. (pp. 1286-1304). Berlin: de Gruyter.

Bowker, L., & Pearson, J. (2002). *Working with specialized language. A practical guide to using corpora*. London: Routledge.

Biel, Ł. (2010). Corpus-based studies of legal language for translation purposes: *Methodological and practical potential. Reconceptualizing LSP. Online proceedings of the XVII European LSP symposium*. Retrieved 30 November 2021 from https://www.asb.dk/fileadmin/www.asb.dk/isek/biel.pdf.

Biel, Ł. (2018). Corpora in institutional legal translation: Small steps and the big picture. In F. P. Ramos (Ed.), *Institutional translation for international governance: Enhancing quality in multilingual legal communication* (pp. 25-36). London: Bloomsbury.

Čmejrková, S. (1996). Academic writing in Czech and English. In E. Ventola & A. Mauranen (Eds.), *Academic writing: Intercultural and textual issues* (pp. 137-152). Amsterdam: John Benjamins.

EAGLES (1996). Preliminary Recommendations on Text Typology. EAGLES Document EAG–TCWG–TTYP/P. Retrieved 10 November 2021 from http://www.ilc.cnr.it/EAGLES/texttyp/texttyp.html

Izquierdo, I. G., & Resurrecció, V. M. (2002). Translating into textual genres. *Linguistica Antwerpiensia*, 1, 135-143.

Hatim, B., & Mason, I. (1990). *Discourse and the translator*. London: Longman.

James, C. (1989). Genre analysis and the translator. *Target*, *1*(1), 29-41. https://doi.org/101075/target.l.l.04jam

Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). Sketch engine. In G. Williams & S. Vessier (Eds.), *Proceedings of the 11th EURALEX International Congress* (pp. 105-115). Université de Bretagne-Sud, Faculté des lettres et des sciences humaines.

Laviosa, S. (2004). Corpus-based translation studies: Where does it come from? Where is it going? *Language Matters: Studies in the Languages of Africa*, *35*(1), 6-27.

Lee, D. Y. W. (2001). Genres, registers, text types, domains and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning and Technology*, *5*, 37-72.

Leech, G. (2007). New resources, or just better old ones? The holy grail of representatives. In M. Hundt, N. Nesselhauf & C. Biewer (Eds.), *Corpus linguistics and the Web* (pp. 133-149). Amsterdam & New York: Rodopi.

Malamatidou, S. (2018). *Corpus triangulation – Combining data and methods in corpus-based translation studies*. New York: Routledge.

Reiss, C., & H. J. Vermeer. (2013). *Towards a general theory of translational action: Skopos theory explained*. London: Routledge.

Schäffner, C. (2000). The role of genre for translation. In A. Trosborg (Ed.), *Analysing professional genres* (pp. 210-224). Amsterdam: John Benjamins.

Snell-Hornby, M. (2010). The turns of translation studies. In Y. Gambier & L. van Doorslaer (Eds.), *handbook of translation studies Vol. 1* (366-370). Amsterdam & Philadelphia: John Benjamins.

Stubbs, M. (2004). Language corpora. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 106-132). Oxford: Blackwell.

Swales, J. M. (1981). *Aspects of article introduction. Aston ESP Research Report no. 1*. Language Study Unit, Birmingham: University of Aston in Birmingham.

Swales, J. M. (1986). A genre-based approach to language across the curriculum. In M. L. Tickoo (Ed.), *Language across the curriculum*. Anthology Series 15 (pp. 10-22). Singapore: SEAMEO Regional Language Centre.

Vermeer, H. J. (1989). Skopos and commission in translational action. In A. Chesterman (Ed.), *Readings in translation theory* (pp. 173-200). Helsinki: Finnlectura.

Zanettin, F. (2000). Parallel corpora in translation studies. Issues in corpus design and analysis. In M. Olohan (Ed.), *Intercultural fault lines* (pp. 105-118). Manchester: St Jerome.