THE JOURNAL OF TEACHING ENGLISH FOR SPECIFIC AND ACADEMIC PURPOSES
Vol. 10, Nº 3, 2022, pp. 547−556

UDC: 159.955.5/.6072-057.875 https://doi.org/10.22190/JTESAP2203547M

Original scientific paper

INTEGRATING CORPUS ANALYSIS TOOLS AND PRACTICES FOR TERMINOLOGY STUDIES IN ESP COURSES

Liudmila Mockienė, Sigita Rackevičienė, Olga Usinskiene

Institute of Humanities, Mykolas Romeris University, Vilnius, Lithuania

Abstract. *This practice illustration paper aims to present the rationale and methodology of using corpus analysis tools and other software to develop students' terminology analysis skills and enhance terminology acquisition at university level. It focuses not only on students of philology, but also on non-linguists within ESP courses. The paper presents the types of assignments for terminological analysis based on corpus analysis methodology as well as skills and competences developed through such activities. Assignments on corpus analysis of terminology provide several important benefits. Besides developing language analysis skills and deepening the knowledge of the specific domain, students also develop their research skills through data collection and analysis. They become researchers who can explore the linguistic phenomena on their own and share their discoveries. They develop the skills of working with different data sources, selection of relevant materials for the research, extracting, interpreting, and managing terminological data. It is hoped that the activities and practices presented in the paper will inspire ESP and EFL teachers to integrate the use of digital language technologies into terminology teaching strategies.*

Key words*: terminology, corpus analysis, teaching specialised vocabulary, vocabulary acquisition, Data-Driven Learning (DDL)*

1. INTRODUCTION

The aim of this practice illustration paper is to introduce terminology analysis methodology applied in the bachelor's degree programme in philology *English for Specific Purposes and the Second Foreign Language (ESPSFL)* and other faculties' bachelor's degree programmes within the courses of *English for Specific Purposes (ESP)* at Mykolas Romeris University (MRU) in Lithuania. The article places great emphasis on the types of assignments completed by students and competences developed in study courses that involve terminology analysis or activities related to terminology extraction and management, as well as corpus analysis on terminology conducted by MRU students. The authors will also illustrate how these activities can be integrated into ESP courses for non-linguists.

As the ESPSFL is a bachelor study programme in philology, students have several theoretical and practical courses which provide them with general lexical and terminological knowledge and linguistic analysis skills, such as Introduction to Linguistics, Corpus-Driven

Submitted October 15th, 2022, accepted for publication November 19th, 2022
Corresponding author: Liudmila Mockienė, Institute of Humanities, Mykolas Romeris University, Vilnius, Lithuania
E-mail: liudmila@mruni.eu

Lexicology, and Computational Terminology Research and Management. However, as the title of the programme suggests their curriculum places great emphasis on ESP, thus students have four courses of English for Specific Purposes and Communication in the areas of Business, Public Administration, Law, and Finance. These ESP courses aim to develop students' skills necessary for successful communication in different professional fields. The objectives of the courses are to provide knowledge of the specific features (distinctive linguistic and stylistic characteristics) of different ESP areas and to develop an overall ability to use ESP flexibly in oral and written intercultural communication. More specifically the learning outcomes of the courses include development of students' skills and abilities to collect data on ESP and carry out their analysis via application of various research methods, work with various information sources, evaluate critically and use practically subject related information, analyse ESP discourse in terms of content and linguistics as well as the ability to organise study and professional activities autonomously, and apply practical and research skills. The skills and abilities that are developed in these ESP courses are universal to any ESP course taught to non-linguists as well.

Achievement of these learning outcomes to a great extent relies on mastery and usage of specific lexical items that are characteristic of a particular ESP field, i.e., specific terminology which is the core of the courses and is developed by means of a variety of discourse analysis assignments and vocabulary acquisition techniques based on application of corpus analysis and other digital tools which will be discussed in more detail.

2. Rationale of Integrating Corpus Analysis and Other Digital Tools into Teaching/Learning Vocabulary

Special vocabulary, i.e., terminology, is the core of LSP as this particular linguistic aspect primarily distinguishes various fields and even specific domains. Acquisition of special vocabulary is tremendously important for developing receptive and productive skills, especially reading comprehension of professional literature and communication with field professionals. Thus, terminology analysis, including the meaning, usage and other linguistic aspects of a term, should be the focus of teaching LSP at more advanced levels.

Corpus linguistics has been used in language research as well as teaching and learning for several decades. As Sinclair (1996) suggests, 'A corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language', which means that such collection of texts or text samples represents the authentic language and can be applied for analysis of language phenomena. It seems that corpus analysis has been used more extensively by linguists for all kinds of language research, translation studies, contrastive analysis, lexicographic activities and the like than in the language classroom.

The idea of introducing corpus analysis for language learning and the concept of data-driven learning (DDL) was first discussed by Johns (1991a, 1991b). DDL has been successfully applied in the language classroom for studying various language aspects, including vocabulary, by a number of scholars (Gabrielatos, 2005; Anderson & Corbett, 2009; Hubbard, 2009; Reppen, 2010; Bennett, 2010; Flowerdew, 2012; Boulton, 2012a; Thomas, 2016; Friginal, 2018; and many others).

One of the assumptions of the DDL approach is that learners "discover" the language and thus 'the language-learner is also, essentially, a research worker whose learning

needs to be driven by access to linguistic data - hence the term "data-driven learning" (DDL) to describe the approach' (Johns, 1991a, p. 2). The role of the teacher in this approach shifts significantly from the conveyer of information to 'the director and coordinator of student-initiated research' (Johns, 1991a, p. 3).

DDL approach has been applied not only in foreign language learning, translator training and philology studies, but also in ESP/EFL. Some scholars have focused their research on application of corpus analysis methodology and tools in ESP courses in general, including needs analysis, syllabus design, genre analysis, ESP writing, etc. (Boulton, 2012b; Nesi, 2012; Boulton, 2016; Bercuci & Chitez, 2019). Meanwhile, some researchers have analysed application of corpus analysis in teaching special lexis within the ESP context specifically (Hou, 2014; Fuentes, 2015; Celiešienė & Juzeleniene, 2020; Malakhovskaja et al., 2021; Rudneva, 2021; Đurović & Bauk (2022).

Corpus software provides a great range of tools which can be applied to analyse special lexis, such as frequency wordlists for establishing the dominant terms within a particular domain, N-Grams for extraction of multiword lexical units, concordance lines for analysis of the use of lexical items and examples, keywords in context (KWIC), collocates, keywords for term extraction, etc.

Corpus analysis allows to develop not only linguistics skills of students, but also a number of other transferable skills and competences, such ability to work in a team, collaborative learning, time-management, people-management, information-management and other skills which are indispensable to students' future career.

3. Practical Application of the Corpus Analysis Methodology and Other Digital Tools in Teaching Terminology

3.1. Introducing the main principles of corpus selection and compilation

To be able to extract relevant and reliable high-quality linguistic data from a corpus students need to be able either to select a relevant corpus that is best representative of the analysed field or domain, or to compile an ad hoc corpus themselves. Both steps involve introduction of certain background knowledge about the main parameters of a corpus and principles of corpus compilation.

The quality of the extracted data and the validity of the research results often depend on such parameters of corpus design as *size* (how large a corpus is), *content* (what types of and how many texts or text samples are included), *representativeness*, which is achieved through *balancing* (what types of genres and in what proportions are included) and *sampling techniques* (does the corpus include whole texts or text fragments) (Sinclair 1991, 1996; Biber, Conrad & Reppen 1998; Kennedy 1998; McEnery & Wilson 2001; McEnery & Gabrielatos 2006; Meyer 2002; Zanettin 2002a, 2000b; Zanettin 2011; Baker 2006; Rundell 2008; Anthony 2009; King 2009; Schäfer, Barbaresi & Bildhauer 2013; Losey-Leon 2015). Other parameters that can also be considered are *the number of texts*, *period coverage* (or time frame), *authorship* (e.g., material produced by native and non-native speakers), and *the source* of the material (open-access, restricted, copyright, etc.). Each of the parameters has been discussed in detail by Valūnaitė Oleškevičienė et al. (2021). The general perception of scholars is that the larger the corpus the better.

The authors' personal experience shows that availability of ready-made special corpora and access to them is rather limited and depends to a great extent on the field or

domain to be explored as well as funding possibilities for acquiring permission to reuse copyright-protected materials. Open access special corpora are usually scarce due to copyright issues. In most cases scholars can use copyright material for research purposes free of charge; however, they cannot publish and share corpora based on such material. Thus, special corpora are usually created for each research project ad hoc.

Depending on the IT literacy and abilities of students and the time available for designing their ad hoc corpora, they can be assigned with the task of selecting the material and compiling the ad hoc corpus completely independently or they might need to be provided with certain degree of guidance from their instructor.

The current corpus building technologies are ample, which makes the process of compiling a corpus fast and easy. Berberich K. and Kleiber I. (2022) have compiled an extensive list of tools for corpus linguistics with descriptions, compatible platforms, links and fees indicated, which is constantly updated by the users and is accessible at https://corpus-analysis.com/. Most of them are completely free of charge, such as *AntConc* (Anthony 2022), some require a fee, while some provide a limited free trial period, such as *Sketch Engine*. What makes introduction of these tools easy for the instructor is that there are numerous detailed video tutorials on the official websites of the software or YouTube starting from the simplest steps such as installing the software to the most complicated tasks such as extraction of specific linguistic data and performing contrastive analysis of specific language features.

3.2. Terminology extraction

After students have explored the relevant domain and selected domain specific texts or text samples to build their ad hoc corpora, they can proceed to the next step of corpus analysis, which is data extraction for further quantitative and qualitative lexical analysis. Students at MRU have access to Sketch Engine as the university covers the annual subscription to this software. However, the authors have successfully used an alternative tool which is free of charge and has similar functionalities – AntConc.

The most basic corpus-driven analysis assignment is extraction of ***most frequent (dominant) lexical items*** from the corpus. To perform this task successfully, students need to be introduced with some theoretical principles of identifying dominant words in a corpus such as the concepts of raw (absolute) and relative frequency, distribution (proportion of texts the word appears in), the difference between a word and a lemma in corpus linguistics, notion words versus function words, parts of speech, the difference between a term and a general language word, etc.

First, students have to extract ten to twenty most frequent single-word terms (the instructor can indicate the number to students) and manage the data by means of several software tools (e.g., tool ***Wordlist*** (available both in Sketch Engine and AntConc) and various functions of MS Excel). As most terms are nouns, students usually extract TOP10 to TOP20 nouns specific to the domain and their statistical metadata (frequency and distribution) from the corpus, and upload and systematize the data in MS Excel file. Some scholars argue that adjectives and verbs can also be considered terms and thus are included in glossaries and dictionaries. If frequency and distribution of domain specific adjectives and verbs is significant, students should include them into their wordlists as well.

In addition, students can extract keywords of their corpus using the tool ***Keywords*** (available both in Sketch Engine and AntConc). This tool compares the ad hoc corpus

with a reference corpus in terms of identifying unique lexical units characteristic of the ad hoc corpus.

As special lexis consists of not only single-word terms but also multi-word units, students can extract them by means of the tool ***N-Gram*** (available both in Sketch Engine and AntConc) by choosing the span from two to four or five words. Students can choose longer N-Grams, however, there is a general tendency of the larger the span, the lower the frequency.

Students can compare the results obtained from their ad hoc corpus with the frequency of the extracted terms in general corpora like the British National Corpus (BNC) or the Corpus of Contemporary American English (COCA). The narrower the domain, the more specific the extracted terms tend to be, thus their frequency in general language corpora is much lower than in special corpora.

Further qualitative analysis of terms, such as collocations, syntactic patterns, derivatives, synonyms, examples, and definitions, requires more input, thus students are advised to choose only several terms (single-word or multi-word) from their lists depending on the number of the group working on a project.

When choosing terms for a more detailed qualitative analysis students or the instructor do not have to rely on frequency alone, which directly depends on the representativeness of the compiled corpus. Terms that designate central concepts of the domain are not necessarily the most frequently used, thus they should not be overlooked for the sake of purely statistical parameters.

3.3. Collocational analysis of terms

To be able to analyse collocates (natural expressions, fixed repetitive patterns that words appear in) of the terms students can use Sketch Engine tools ***Word Sketch*** and ***Collocations*** or AntConc tool ***Collocates***. The Collocations and Collocates tools allow students to choose the span of collocates to the left, i.e., words that modify the term, and to the right, i.e., words that are modified by the analysed term. To get a larger pattern of the term in use students can also use the tool ***Concordance*** (available both in Sketch Engine and AntConc). In addition, the Sketch Engine tool ***Word Sketch*** provides lists of collocations categorised according to their grammatical relations and frequency and typicality scores. These tools can be used both for analysis of collocations of single-word terms and identification of multi-word terms.

After analysing the lists of collocates in the corpus, students make a list of most useful collocations the term appears in, including modifying adjectives, nouns and verbs. Students can also note down any specific syntactic patterns that the term is used in, especially the ones with prepositions.

Collocational analysis can be used to differentiate polysemous, homonymous and synonymous terms. Ideally, a term has to be monosemous and a concept should be designated by one term to ensure consistency and avoid ambiguity. However, due to different processes of terminology development, such as borrowing, terminologization, transterminologization, jargonization and the like, some terms can have several meanings and several terms can refer to the same concept, e.g., in legal terminology *disposition* is a polysemous term as it has four meanings (see Chromá 2011), while in accounting terminology *revenue, income, receipts,* and *profit* are synonyms (see Vogel 2008).

The empirical analysis of a polysemous term starts with extraction of collocations of a polysemous term by using *Word Sketch* and *Concordance* tools in Sketch Engine (or AntConc) and their classification according to the meanings of the polysemous term in them in MS Excel. Students are asked to distinguish the meanings of the polysemous term, describe them with a short definition and illustrate them with three to five typical collocations. Students visualise the analysis results in a mind map drafted by using MindMup (or another mind mapping app). Next, students analyse the meanings of the same polysemous term in a dictionary and compare them with the ones established in the corpus by answering the provided questions. In conclusion, students describe what primary and secondary (metaphoric/metonymic) meanings they have established and in which of them the polysemous word in used most often in the corpus; they also discuss the results of the comparative corpus and dictionary analysis.

The empirical analysis of a synonym set (or pair) starts with establishing synonyms: students choose a term, establish its closest synonyms using Sketch Engine tool *Thesaurus* and online English thesauri and select one or more of the synonyms for further analysis. Next, students establish which of the synonyms has higher frequency and distribution in the corpus and present the results in a chart. Subsequently, students extract collocations of synonyms by using *Word Sketch Difference*, *Word Sketch* and *Concordance* tools and systematise them in MS Excel. Students are asked to select at least 10 most typical collocations for each synonym and group them according to similarities and differences in the meanings of the synonyms. Students visualise the analysis results in a mind map by using MindMup (or another mind mapping app). Next, students analyse the meanings of the chosen synonyms in a specialised dictionary and compare them with the findings of their corpus analysis by answering the provided questions. In conclusion, students describe the differences between the synonyms established in the corpus (their frequency differences and collocational differences) and discuss the results of the comparative corpus and dictionary analysis.

These assignments develop further lexical data analysis skills, as well as enable students to obtain essential knowledge on polysemy and synonymy.

3.4. Analysis of derivatives of single-word terms

Terminology of most domains constantly develops, thus new terms and their derivatives are introduced and used by professional communities much faster than they are fixed in dictionaries. Therefore, corpus analysis of an up-to-date ad hoc corpus can provide a greater variety of derivatives of terms than a dictionary.

To be able to extract as many derivatives of a term as possible students enter the root of a term with asterisks on both sides, e.g., *govern*, in *Concordance* of a corpus and analyse the word-forms in detail to note down as many prefixal and suffixal derivatives of the term and its compounds as possible. This method allows to extract not only suffixal derivatives, but also prefixal derivatives and compounds with the term as the core rather than the modifier.

Next, students can categorize the term derivatives according to several aspects: the part of speech and the means of term-formation. They also can establish the relations between the term derivatives, that is the forms that were used to create other words, e.g., *government* is a derivative of *govern* formed by means of suffix *-ment*, while *governmental* is a derivative of *government* formed by means of suffix *-al* and *antigovernmental* is derived from *governmental* by means of prefixation – *anti-*. Examples of compounds are *self-governing* and

government-funded. Students can visualise the derivational analysis results in a mind map by using MindMup (or another mind mapping app).

Such linguistic analysis draws students' attention to certain word-formation patterns and trends in developing terminology of a particular domain, which is important for their future professional life as they will have to use and possibly create new terminology as specialists of their fields.

3.5. Management of the extracted terminological data

Modern technologies offer several options to systematize and manage terminological data. One of them that is widely used by MRU students and professors is *Terminologue*, a free of charge open-source terminology management tool, which can be either downloaded or installed as one's own instance or can be used online as a cloud-based instance for group work.

Terminologue allows several students to work on the same terminological project, share the created termbases with other users and make them available to the public. It is very easy to operate and there are detailed instructions with screenshots for beginners on its welcome page.

Entries created in *Terminologue* termbases cover a wide range of terminological data, such as indicating domains and subdomains the term belongs to, definitions, examples, synonyms, translations, collocations, notes, etc. Depending on students' abilities and the aim of the terminological project they work on, they can choose to fill in all or only some of the possible fields in each entry.

The source of the material used to compile the ad hoc corpus (such as textbooks, scientific articles, manuals, legal acts, product descriptions, etc.) determines whether students can extract definitions of terms from the corpus. Alternatively, they can use specialised dictionaries and other terminology databases, such as IATE (the EU's interinstitutional terminology database) to define terms. *Terminologue* enables users to indicate the source of the definition.

4. Final Remarks and Conclusions

Assignments on corpus analysis of terminology provide several important benefits. Students develop their research skills through data collection and analysis and at the same time gain more knowledge on the linguistic aspects of their fields. Students become researchers who can explore the linguistic phenomena on their own and share their discoveries. They develop the skills of working with different data sources, selection of relevant materials for the research, extracting, interpreting, and managing terminological data. The added value of such terminological research is deepening the specialized knowledge of the analysed domains and greater awareness of how terms of a particular domain are used in context and formed.

Corpus analysis methodology and other digital tools can be successfully applied in ESP courses taught to both linguists and non-linguists. Depending on the level of students and their linguistic and IT background, they can require a certain degree of input regarding the major principles of corpus analysis and linguistic analysis.

Nonetheless, corpus analysis of terminology covers a great range of activities of various complexity, which provides incredible flexibility in choosing the relevant tasks for a particular group of students.

Although corpus analysis has not been widely applied in ESP courses yet, the authors hope that the activities and practices presented in this practice illustration paper will inspire ESP and EFL teachers to integrate the use of corpus analysis into terminology teaching strategies.

REFERENCES

Anderson, W., Corbett, J. 2009. *Exploring English with online corpora.* London: Palgrave Macmillan.

Anthony, L. 2009. "Issues in the Design and Development of Software Tools for Corpus Studies: The Case for Collaboration". In *Contemporary Corpus Linguistics*, edited by P. Baker, 87-104. London, UK: Continuum Press.

Baker, P. 2006. *Using Corpora in Discourse Analysis.* London and New York: Continuum.

Bennett, G. 2010. *Using corpora in the language learning classroom: Corpus linguistics for teachers*. Michigan: University of Michigan Press.

Berberich, K., Kleiber, I. 2022. "Tools for Corpus Linguistics." Available from https://corpus-analysis.com/.

Bercuci, L., Chitez, M. 2019. "A Corpus Analysis of Argumentative Structures in ESP Writing." *International Online Journal of Education and Teaching*, 6(4): 733-747.

Biber, D., Conrad S., Reppen, R. 1998. *Corpus Linguistics: Investigating Language Structure and Use.* Cambridge: Cambridge University Press.

Boulton, A. 2012a. "What Data for Data-Driven Learning?" *EuroCALL Review* 20(1): 23-27.

Boulton, A. 2012b. "Corpus Consultation for ESP: A Review of Empirical Research." *Corpus-informed Research and Learning in ESP: Issues and Applications*: 261-291.

Boulton, A. 2016. "Integrating Corpus Tools and Techniques in ESP Courses". *ASp. la revue du GERAS* (69): 113-137.

Celiešienė, V., Juzeleniene, S. 2020. "Semantic Motivation of the Terminologized Lexis in the Field of Digital Technologies." *Journal of Teaching English for Specific and Academic Purposes*: 031-042.

Chromá, M. 2011. "Synonymy and Polysemy in Legal Terminology and their Applications to Bilingual and Bijural Translation." *Research in language* 9(1): 31-50.

Đurović, Z., Bauk, S. 2022. "Testing Corpus Linguistics Methods in ESP Vocabulary Teaching." *Journal of Teaching English for Specific and Academic Purposes* 10(2): 319-337.

Flowerdew, L. 2012. "Corpora and Language Education. In *International Journal of Corpus Linguistics*. Basingstoke: Palgrave Macmillan: 147-155.

Friginal, E. 2009. *Corpus Linguistics for English Teachers: Tools, online Resources, and Classroom Activities.* Abingdon/New York: Routledge.

Fuentes, A. C. 2015. "Lexical Acquisition in ESP via Corpus Tools: Two Case Studies." *Scripta Manent*, 3(1): 21-34.

Hou, H. I. 2014. "Teaching Specialized Vocabulary by Integrating a Corpus-Based Approach: Implications for ESP Course Design at the University Level." *English Language Teaching*, 7(5): 26-37.

Hubbard, P. 2009. *Computer-Assisted Language Learning*. Routledge.

Johns, T. 1991a. "Should You Be Persuaded: Two Examples of Data-Driven Learning". In *Classroom Concordancing,* edited by Johns, T., King, P. Birmingham: ELR, Vol. 4: 1-16.

Johns, T. 1991b. “From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-Driven Learning.” *ELR Journal (Birmingham)*, Vol. 4: 27-45.

Kennedy, G. 1998. *An Introduction to Corpus Linguistics.* London: Longman.

King, B. 2009. “Building and Analysing Corpora of Computer-Mediated Communication”. In *Contemporary Corpus Linguistics*, edited by Baker, P. London and New York: Continuum. Continuum: 301-320.

Losey-Leon, M. A. 2015. “Corpus Design and Compilation Process for the Preparation of a Bilingual Glossary (English-Spanish) in the Logistics and Maritime Transport Field: LogisTRANS”. *Procedia - Social and Behavioral Sciences*, Elsevier Ltd., Vol. 173: 293-299.

Malakhovskaya, M., Beliaeva, L., Kamshilova, O. 2021. “Teaching Noun-Phrase Composition in EAP/ESP Context: A Corpus-assisted Approach to Overcome a Didactic Gap”. *Journal of Teaching English for Specific and Academic Purposes*, 9(2): 257-266.

McEnery, T., Wilson, A. 2001. *Corpus Linguistics*. 2nd ed. Edinburgh: Edinburgh University Press.

McEnery, T., Gabrielatos, C. 2006. “English Corpus Linguistics.” In *the Handbook of English Linguistics*, edited by Aarts, B. & McMahon, A. Oxford: Blackwell 33-71.

Meyer, C. F. 2002. *English Corpus Linguistics.* Cambridge, UK: Cambridge University Press.

Nesi, H. 2012. “ESP and Corpus Studies.” In *the Handbook of English for Specific Purposes*: 407-426.

Reppen, R. 2010. *Using Corpora in the Classroom.* Cambridge: Cambridge University Press.

Rudneva, M. 2021. “Corpus-Driven ESP Pedagogy: A Preliminary Case Study.” *Journal of Teaching English for Specific and Academic Purposes*: 241-248.

Rundell, M. 2008. “The Corpus Revolution Revisited.” *English Today* 24 (1): 23-27.

Schäfer, R., Barbaresi, A. Bildhauer F. 2013. “The Good, the Bad, and the Hazy: Design Decisions in Web Corpus Construction.” In *Proceedings of the 8th Web as Corpus Workshop*, edited by Evert, S., Stemle, E. and Rayson P.: 7-15.

Sinclair, J. M. 1991. *Corpus, Concordance Collocation*. Oxford: Oxford University Press.

Sinclair, J. 1996. “Preliminary Recommendations on Corpus Typology.” *EAGLES Document TCWG-CTYP/P*.

Thomas, J. 2016. *Discovering English with Sketch Engine. A Corpus-Based Approach to Language Exploration.* 2nd ed. Brno: Versatile.

Valūnaitė Oleškevičienė, G., Mockienė, L., Stojković, N. 2021. “Corpus Analysis for Language Studies at the University Level.” *Journal of Teaching English for Specific and Academic Purposes*. Cambridge Scholars Publishing, 9(4): 785−786.

Vogel, R. 2008. “Synonymy and Polysemy in Accounting Terminology: Fighting to Avoid Inaccuracy.” *SKASE Journal of Translation and Interpretation*, 3(1): 91-102.

Zanettin, F. 2002a. “CEXI: Designing an English Italian Translational Corpus.” *Language and Computers.* 329–43.

Zanettin, F. 2002b. “DIY Corpora: The WWW and the Translator.” In *Training the Language Services Provider for the New Millennium, edited by Belinda Maia, Jonathan Haller and Margherita Urlrych*, 239–48. Porto: Facultade de Letras, Universidade do Porto.

Zanettin, F. 2011. “Translation and Corpus Design.” *A Journal of Professional Communication. Synaps* 26: 14–23.

SOFTWARE

AntConc. Anthony, L. 2022. Version 4.1.4 [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software.
Microsoft Excel. Microsoft Corporation. https://office.microsoft.com/excel/
MindMup. Sauf Pompiers Ltd. https://www.mindmup.com/
Sketch Engine. Lexical Computing. http://www.sketchengine.eu
Terminologue. Gaois research group at Dublin City University. 2019. The loud-based instance Terminologue.org: https://www.terminologue.org/