Using Corpus Tools for Stylistic Analysis in Translator Training

Çevirmen Eğitiminde Derlem Araçlarının Biçembilimsel Çözümleme Amacıyla Kullanılması

Özgür Şen Bartan[1] (ID)

ABSTRACT

This study presents the findings from two case studies, in which literary texts and their translation corpora were used in the training of BA and MA level students in a translation and interpreting department. It suggests methods of using corpus-based resources in the BA and MA literary translation context. During each case study, the students were guided to use corpus tools such as Lextutor and Voyant to analyse the style of the author and/or the translator concentrating on different aspects of translations of style which encompass dimensions such as lexical category, grammatical category, figurative language and cohesion, and to compare the parallel corpora. Moreover, the study provides an evaluation of the materials with a user perspective survey. In total, 39 students participated in the study, and 18 students shared their views on using corpus tools. As a conclusion, the findings suggest that, in general, the corpus tools can be used to analyse literary texts and their translations focusing on their styles in translator training. The views of the participants are mostly positive, and they find the process helpful in both analysing literary texts and their personal translation works.

Keywords: Corpus-based studies, corpus stylistics, corpus tools, literature for translators, translator training

[1]Assist. Prof. Kırıkkale University, Faculty of Arts and Sciences, Department of Western Languages and Literatures, Department of Translation and Interpretation (English), Kırıkkale, Türkiye

ORCID: Ö.Ş.B. 0000-0002-7351-7858

Corresponding author/Sorumlu yazar:
Özgür Şen Bartan,
Kırıkkale University, Faculty of Arts and Sciences, Department of Translation and Interpretation, Dudullu Campus, Istanbul, Turkiye
E-mail: ozgursen1@yahoo.com

ÖZ

Bu çalışma, mütercim ve tercümanlık bölümü lisans ve yüksek lisans öğrencilerinin eğitiminde yazınsal metinlerin ve çeviri derlemlerinin kullanıldığı iki vaka çalışmasından elde edilen bulguları sunmaktadır. Lisans ve yüksek lisans düzeylerinde yazınsal çeviri bağlamında derlem tabanlı kaynaklardan yararlanma yöntemleri önerilmektedir. Her bir çalışma sırasında, öğrenciler Lextutor ve Voyant adlı derlem araçlarını kullanarak yazarın ve/veya çevirmenin biçeminin alt başlıkları olan sözcüksel kategori, dilbilgisel kategori, değişmeceli dil ve bağdaşıklık gibi farklı yönlerine odaklanarak derlemleri karşılaştırıp çözümlemişlerdir. Çalışmada ayrıca kullanıcı perspektifli kısa bir sormaca ile kullanılan araçların değerlendirmesi sunulmaktadır. Toplam olarak 39 üniversite öğrencisi çalışmaya katılmıştır ve 18 öğrencinin derlem araçlarıyla ilgili görüşleri elde edilmiştir. Sonuç olarak, bulgular, genel olarak, derlem araçlarının çevirmen eğitiminde yazınsal metinleri ve çevirilerini biçemlerine odaklanarak çözümlemek için kullanılabileceğini göstermektedir. Katılımcı görüşleri çoğunlukla olumludur ve araçların kullanımını hem yazınsal metinlerin çözümlemesinde hem de kişisel çeviri çalışmalarında yararlı bulmaktadırlar.

Anahtar kelimeler: Derlem temelli çalışmalar, derlem biçembilimi, derlem araçları, çevirmenler için edebiyat, çevirmen eğitimi

Introduction

It is becoming increasingly difficult to ignore the effect of new technologies in language studies, and the past decade has seen the rapid development of corpus tools in many fields such as linguistics, foreign language teaching and translation and interpreting studies (TIS). These rapid changes are having a serious effect on the curriculum development of TIS. Hansen and Shlesinger (2007, p.96) recommended that new technologies can be used in order to create more self-study materials. There are many studies (Baker, 1993; Bale, 2015; Corpas Pastor, 2018; Nebot, 2008; Olohan, 2003; Olohan, 2004; Pearson, 2003) suggesting the use of corpus tools in translator and interpreter training.

Literary translation is regarded as one of the most difficult translation fields, and it needs a preparation phase and awareness of the author's style. Nord (1991, p.1), similar to many translation theorists, proposed translation-oriented textual analysis, and within the framework of this functional concept, translators "[…] can then choose the translation strategies suitable for the intended purpose of the particular translation they are working on". Also, İnce (2019, p.128), who is a well-known and distinguished literary translator in Turkey, states that it is not enough to know the languages in order to translate, and also she insists that 'to be good translators, they need to be good at reading and writing, and they need to know how to read and write a text'. Through Akalın (2022, p.188), this perspective from practice is associated with the idea (See Kurultay, 1995; Nord, 1991) from translation theory that translators need to analyse the source text before translation. It is clearly seen that theory and practice in translation are aligned in terms of an analysis of the source text before translation, and also it is clearly seen that before translation, translators need to have a preparation phase. This phase should include systematic steps, and this is specifically crucial for novice and/or candidate translators who may have different levels of translation skills and a varied level of knowledge in literature.

In the context of literary translation, translation and interpreting programs offer courses in literature and/or stylistics for translators in order to build background knowledge for literary translation. They aim to prepare candidate translators to explore various dynamics of fiction and poetry. According to Kılıçkaya and Ergil (2023, p.255), literature courses "help students acquire and develop linguistic, analytical, critical, and academic competences in general and critical thinking, critical reading, analytical thinking, cultural awareness competences as well as linguistic and textual competence in particular". The researchers posed a question concerning "how students will learn the methods required to decode the text and make expert critical judgments about the text" and proposed that "the most appropriate key for this appears to be stylistic analysis" (p.255).

The purpose of this study is to explore how corpus tools can be used to analyse literary texts and their translations, and to implement methods of using corpus tools in literary translation context for the BA and MA level students of translation and interpreting in a university in Turkey. It suggests methods of using corpus tools in the BA and MA literary translation context.

During each implementation, students were guided in using corpus tools such as Lextutor and Voyant to analyse the style of the author and/or the translator concentrating on different aspects/categories of translations of style such as the lexical category, grammatical category, figures of speech and cohesion, and lastly their views on using corpus tools were collected through an online survey.

Accordingly, this paper seeks to address the following questions:

1. How can corpus tools be used to analyse literary texts and their translations focusing on their style in translator training?
2. What are the views of students who use corpus tools to analyse the style of literary texts and their translations?

The article has been organised in the following way. It first gives a brief overview of the background to the corpus-based studies of translation and stylistic analysis categories. The second part presents the methodology of the study which consists of the procedures, setting, participants, and the corpus tools. The third part conveys the findings, finally, followed by a conclusion.

Literature Review

Corpus-based studies of translation

Reppen (2010) defines corpus as "a large, principled collection of naturally occurring language texts (written or spoken) stored electronically" (p.3). It is a collection of spoken or written texts that are compiled generally in an electronic form. Some examples of corpus are BNC (British National Corpus), COCA (Corpus of Contemporary American English), MICASE (Michigan Corpus of Academic Spoken English), TDB (Turkish Discourse Bank), PDTB (Penn Discourse Tree Bank), TED-MDB (TED Multilingual Discourse Bank), TEC (Translated English Corpus). There are different types of corpus such as monolingual, parallel, and comparable.

Hansen and Shlesinger (2007, p.96) recommended that new technologies can be used in order to create more self-study materials. There are many studies (Bale, 2015; Corpas Pastor, 2018; Nebot, 2008; Olohan, 2003; Olohan, 2004; Pearson, 2003) that suggest using corpus in translator and interpreter training. In short, a considerable amount of literature has been published on corpus-based translation studies.

Hu (2016) outlines corpus-based studies of translation as features of translation (translation universals), style ((in)visibility), translation norms, translation practice, translation teaching, and interpreting studies. Corpus-based translation studies was initiated by Baker (1993) and defined as "the branch of the discipline that uses corpora of original and/or translated text for the empirical study of the product and process of translation, the elaboration of theoretical constructs and the training of translators" by Laviosa (2004, p.45). According to Laviosa

(2004), together with corpus linguistics, corpus-based translation studies "generated a coherent, composite and rich paradigm that addresses a variety of issues pertaining to theory, description and the practice of translation" (Laviosa, 1998).

Mikhailov (2022, p.224) studied to what extent corpora are present in university translator training programmes, and found that at least in EU countries, it is becoming a part of curricula of TS programmes. The researcher criticized these programmes for not including the instruction of "compiling Do-It-Yourself corpora" which is widely used as a research instrument rather than as a tool, and is used instead of practical translation work. Güner (2018, p. iv) also concluded that the corpus model which was designed by the researcher for translation studies "could help increase the quality of translations" and in general, "corpus-based approaches may contribute to both theoretical and applied Translation Studies". The researcher also suggested that machine translation systems need to be integrated into the corpus model.

Corpus linguistic methodologies have been used in stylistic analysis, and in literature, researchers have shown an interest in studying corpus stylistics (Mahlberg, 2013; McIntyre, 2013). Corpus stylistics is defined as "the application of corpus methods to the analysis of literary texts by relating linguistic description with critical interpretation." (Mahlberg, 2013, p. 1282). Corpus stylistic studies mostly focus on "lexical patterns", and they use methods such as concordancing, "the retrieval of clusters (also called n-grams), that is, repeated sequences of words such as "I don't know what", or "the generation of key words" (Mahlberg & Wiegand, 2020, p.312). Key words are the repeated words that were used frequently in a text. Additionally, Mastropierro and Mahlberg (2017, p.78) used key words to identify cohesive networks in a novel and discussed "how these networks contribute to the construction of literary meanings in the text". According to the researchers, key words can reflect the themes of a text and provide insights into what a text is about. They showed the potential of a corpus approach to studying translated cohesion, and recommended applying models and methods from corpus linguistics to corpus stylistic analysis (p.78).

In this field, Baker (2000) carried out the first study in a corpus-based study of two translators' style. The researcher compared the styles of two British translators Peter Bush and Peter Clark. She tried to find out whether translators exhibit their own style in the translation of literary works and whether they have a preference in the use of words, syntactic structures, or punctuations. The two translators' styles have many differences such as type / token ratio, average sentence length, and narrative structure (the past tense and direct speech/indirect speech).

Olohan (2003) compared the translations by Peter Bush and Dorothy S. Blair, and found that Bush has the tendency to use contracted forms more often, while Blair prefers complete forms. Also, Olohan (2003) compared the translations by Lawrence Venuti and revealed that Venuti tends to adopt the first-person perspective and uses academic connectives in novels such as "yet," "nonetheless," "nor," and "thus" and prefers formal and archaic words.

Furthermore, in the literature on translator training, bilingual parallel corpora have been studied (Zanettin, 1994; Nebot, 2003; Nebot, 2008; Pearson, 2003; Zanettin, 2007), and they "expose students to texts translated by professionals (Nebot, 2008, p.244). Figure 1 shows the areas of uses of bilingual parallel corpora in legal translation training.
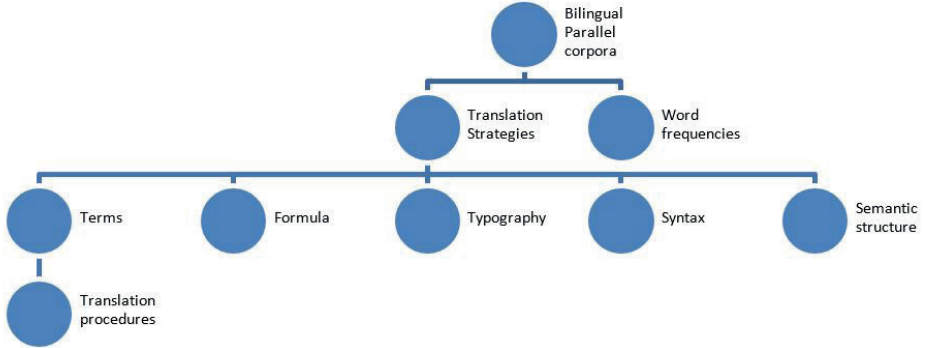

Figure 1. Uses of bilingual parallel corpora in legal translation training (Nebot, 2008, p. 244).

For the purposes of the implementation of this study, the stylistic analysis categories of Leech and Short (2007) were adopted, and the participants of the study were encouraged to use this method of analysis through corpus tools.

Stylistic Analysis

The present study adopts Leech and Short's (2007) stylistic analysis categories during the case studies, therefore, this section will briefly outline the principles of the researchers in analysing the style of a prose text.

According to Leech and Short (2007), every style is unique, and they state in the following extract from their book *Style in Fiction* that there is not a perfect method for analysis of style:

> Every analysis of style, in our terms, is an attempt to find the artistic principles underlying a writer's choice of language. All writers, and for that matter, all texts, have individual qualities. Therefore the features which call themselves to our attention in one text will not necessarily be important in another text by the same or a different author. There is no infallible technique for selecting what is significant. We have to make ourselves newly aware, for each text, of the artistic effect of the whole, and the way linguistic details fit into this whole (p.60).

Although they admitted that no method is "infallible", they found it useful to have a "checklist of features and questions which may or may not be significant in a given text (p.61). Additionally, using all the answers to the questions in the checklist, it is recommended to make a general comment on each category and sub-category. Table 1 summarises the researchers' method of analysis.

Table 1. Stylistic analysis categories of Leech and Short (2007) (see the full version Leech and Short, 2007).

Stylistic Categories	Sub-categories	Questions
Lexical Category	Nouns	Are the nouns abstract or concrete? What kinds of abstract nouns occur (e.g. nouns referring to events, perceptions, processes, moral qualities, social qualities)? What use is made of proper names? Collective nouns?
	Adj.	Are the adjectives frequent? To what kinds of attribute do adjectives refer? Physical? Psychological? Visual? Auditory? Colour? Referential? Emotive? Evaluative? etc. Are adjectives restrictive or non-restrictive? Gradable or non-gradable? Attributive or predicative?
	Verbs	Do the verbs carry an important part of the meaning? Are they stative (referring to states) or dynamic (referring to actions, events, etc.)? Do they 'refer' to movements, physical acts, speech acts, psychological states or activities, perceptions, etc.? Are they transitive, intransitive, linking (intensive), etc.? Are they factive or non-factive?
	Adv.	Do the verbs carry an important part of the meaning? Are they stative (referring to states) or dynamic (referring to actions, events, etc.)? Do they 'refer' to movements, physical acts, speech acts, psychological states or activities, perceptions, etc.? Are they transitive, intransitive, linking (intensive), etc.? Are they factive or non-factive?
Grammatical Category	Sentence types	Does the author use only statements (declarative sentences), or do questions, commands, exclamations or minor sentence types (such as sentences with no verb) also occur in the text? If these other types appear, what is their function?
	Sentence complexity	Do sentences on the whole have a simple or a complex structure? What is the average sentence length (in number of words)? What is the ratio of dependent to independent clauses? Does complexity vary strikingly from one sentence to another? Is complexity mainly due to (i) coordination, (ii) subordination, or (iii) parataxis (juxtaposition of clauses or other equivalent structures)? In what parts of a sentence does complexity tend to occur? For instance, is there any notable occurrence of anticipatory structure (e.g. of complex subjects preceding the verbs, of dependent clauses preceding the subject of a main clause)?
	Clause types	What types of dependent clause are favoured: relative clauses, adverbial clauses, different types of nominal clauses (that-clauses, wh-clauses, etc.)? Are reduced or non-finite clauses commonly used and, if so, of what type are they (infinitive clauses, -ing clauses, -ed clauses, verbless clauses)?
	Clause structure	Is there anything significant about clause elements (e.g. frequency of objects, complements, adverbials; of transitive or intransitive verb constructions)? Are there any unusual orderings (initial adverbials, fronting of object or complement, etc.)? Do special kinds of clause construction occur (such as those with preparatory it or there)?

Figurative Language	Grammatical and lexical	Are there any cases of formal and structural repetition (anaphora, parallelism, etc.) or of mirror-image patterns (chiasmus)? Is the rhetorical effect of these one of antithesis, reinforcement, climax, anti-climax, etc.?
	Phonological schemes	Are there any phonological patterns of rhyme, alliteration, assonance, etc.? Are there any salient rhythmical patterns? Do vowel and consonant sounds pattern or cluster in particular ways? How do these phonological features interact with meaning?
	Tropes	Are there any obvious violations of, or departures from, the linguistic code? For example, are there any neologisms (such as Americanly)? Deviant lexical collocations (such as portentous infants)? Semantic, syntactic, phonological, or graphological deviations? Such deviations (although they can occur in everyday speech and writing) will often be the clue to special interpretations associated with traditional poetic figures of speech such as metaphor, metonymy, synecdoche, paradox and irony. If such tropes occur, what kind of special interpretation is involved (e.g. metaphors can be classified as personifying, animising, concretising, synesthetic, etc.)? Due to its close connection with metaphor, simile may also be considered here. Does the text contain any similes, or similar constructions (e.g. 'as if' constructions)? What dissimilar semantic fields are related through simile?
Cohesion and context	Cohesion	Does the text contain logical or other links between sentences (e.g. coordinating conjunctions, or linking adverbials)? Or does it tend to rely on implicit connections of meaning? What sort of use is made of cross-reference by pronouns (she, it, they, etc.)? by substitute forms (do, so, etc.), or ellipsis? Alternatively, is any use made of elegant variation – the avoidance of repetition by the substitution of a descriptive phrase (as, for example, 'the old lawyer' or 'her uncle' may substitute for the repetition of an earlier 'Mr Jones')? Are meaning connections reinforced by repetition of words and phrases, or by repeatedly using words from the same semantic field?
	Context	Does the writer address the reader directly, or through the words or thoughts of some fictional character? What linguistic clues (e.g. first-person pronouns I, me, my, mine) are there of the addresser–addressee relationship? What attitude does the author imply towards his or her subject? If a character's words or thoughts are represented, is this done by direct quotation (direct speech), or by some other method (e.g. indirect speech, free indirect speech)? Are there significant changes of style according to who is supposedly speaking or thinking the words on the page?

Like many other researchers, Tutaş (2006) used this method of stylistic analysis on literary texts, analysing a short poem by E.E. Cummings and its translation. Having explored a short poem and its translation through Leech and Short's stylistic analysis method, the researcher recommended using this method to "explain and support the readers' intuitive understandings of texts, and their responses to them." (2006, p.169). According to her, the reader can use this 'tool-kit' to support their comments on style. This stylistic analysis method which was suggested by Leech and Short (2007) is adopted for this study to implement it through corpus tools.

Method

For this study, a case study method, which is a method of qualitative research, was conducted to explore how corpus tools can be used to analyse literary texts and their translations in a literary translation context for the BA and MA level students of translation and interpreting in a university in Turkey for 6 weeks of a 14-week course. Additionally, via an online questionnaire, the participants' views were collected. Creswell defines case studies as qualitative and pre-experimental research designs (2007, pp.14,19), and they explore processes, activities, and events (p.21).

This section will present the procedures of the case studies, followed by the setting and the participants of the study since case studies encompass a detailed description of the setting or intervals (Wolcott, 1994). The next sub-section will be on the corpus tools which are the research instruments.

Setting and the participants

Thirty six BA level 2nd year students studying in the Department of English Translation and Interpreting were recruited for this study. They took the course called Literature for Translators in the 2021-2022 academic year Fall Semester. The syllabus of the course also covers the following topics:

- What is literature? Its role, functions and features
- The Importance of literature and creative writing for translators
- Elements of a story: Setting, plot, characters, story openings, how to hook the reader, conversation in the novel, language and style
- What is style? Approaches and methods: Dualism, Monism, Pluralism
- What is Stylistics?
- Stylistic Categories of Leech and Short: Lexical Category
- Stylistic Categories of Leech and Short: Grammatical Category
- Stylistic Categories of Leech and Short: Figurative Language
- Stylistic Categories of Leech and Short: Context and Cohesion
- A sample poem to analyse
- Corpus tools for stylistics analysis and implementations
- Reading a short story and analysing its style
- Reading a children's story and analysing its style
- Reading a play and analysing its style

This course was supported by different resources such as books, articles, and online tools. Additionally, two Edmodo education classrooms were utilized as online education platforms to communicate, to upload resources, and to give assignments, instructions and quizzes during education for both the BA and MA levels. Also for the assignments they were supposed to

complete, all the instructions and messages related to carry out the tasks were sent through Edmodo classrooms. Moreover, they were scaffolded when needed through both direct messages and face to face, since they were totally inexperienced in using corpus tools and Turnitin Plagiarism Software. That's why, the tutorials on how to use the tools were shared through Edmodo as well[1]. All the participants were supposed to become members of Turnitin Plagiarism account of the researcher so as to construct similarity report for their assignments. The tutorials concerning how to submit an assignment for checking in Turnitin were shared through Edmodo classroom as well. Therefore, using Turnitin guided them to generate original work, and encouraged them to build an ethical culture in their educational environment.

Also, the participants (BA) had previously taken translation courses such as Introduction to Translation, Text Translation, Translations of Technical and Scientific Text, Speaking 1&2, Linguistics, Grammar 1&2, Reading, Vocabulary, and Intercultural Communication and History of Humanity.

Moreover, 3 MA level students studying in the Department of English Translation and Interpreting Master's Degree Program were recruited for this study. They took the course called *Stylistics/Literary Translation* in the 2021-2022 academic year Spring Semester. The syllabus of the course called Stylistics/Literary Translation covers the following topics: what is literary translation, problems and solutions of literary translation, translation strategies, translation strategies of metaphor and idiom, cultural elements, literature translation analysis, translation of children's literature and problems and possible solutions of it, what is style, a method of stylistic analysis and its application, what is corpus stylistics and studies in the field, how to create a bilingual translation corpus, how to use corpus tools for stylistic analysis, tasks and/ or research questions on stylistic analysis via corpus tools.

Out of three MA students, one was a graduate of a Translation and Interpreting BA Program, the other two were graduates of English Language Teaching Programs.

In total, as Table 2 shows, 39 students participated in the study and 18 students shared their views (see Table 3) on using corpus tools. The data covering students' views was collected through Google Documents online. The study was approved by the Kırıkkale University Ethics Committee.

Table 2. The participants of the case studies.			
	BA level	**MA level**	**Total**
Male	16	2	13
Female	21	1	26
Total	**37**	**3**	**39**

---

1   The tutorials can be found on these links: https://www.youtube.com/watch?v=LWyMU0fWX58; https://www. lextutor.ca/instructions.htm

Table 3. The participants of the survey.

	BA level	MA level	Total
Male	4	1	5
Female	12	1	13
Total	**16**	**2**	**18**

Procedures

To address the issues outlined above, a corpus-based approach to the training of the students was implemented over a period of six weeks. Table 4 and 5 shows the weekly plan of the implementations.

Table 4. Procedure for the BA level students.

Weeks	Topics
Week 1 (90 minutes)	What is style?
Week 2 (90 minutes)	Stylistic analysis types and Leech and Short (2007)
Week 3 (90 minutes)	Corpus Tools and sample studies
Week 4 (90 minutes)	Harry Potter (ST&TT)sample analysis
Week 5 (90 minutes)	A short story sample analysis (ST)
Week 6 (90 minutes)	**Assignment:** Sample analysis via corpus tools: *The Suspicions of Mr Whicher or The Murder at the Road Hill House* by Kate Summerscale

Table 5. Procedure for the MA level students.

Weeks	Topics
Week 1 (3 hours)	Problems of literary translation: cultural elements, figures of speech etc. / Possible Solutions to the problems of literary translation
Week 2 (3 hours)	Domestication and foreignization strategies
Week 3 (3 hours)	What is style? Stylistic analysis methods and Leech and Short's (2007) stylistic analysis method
Week 4 (3 hours)	Corpus Stylistics and Methods and Tools/Sample studies and analysis methods
Week 5 (3 hours)	Sample analysis of TED Talks (ST&TT) focusing on cohesion, coherence and rhetorical devices
Week 6 (3 hours)	**Assignment:** Sample analysis via corpus tools and Microsoft Excel: *Benim Adım Kırmızı/My name is Red* by Orhan Pamuk, and its English translation by Erdağ Göknar.

Corpora and corpus tools used in this study

Many corpus tools have been developed such as Wordsmith, AntConc, AntPConc, Voyant, Lextutor, LanchBox etc. Lextutor and Voyant corpus tools were chosen to be used for the purposes of this study. Lextutor and Voyant corpus tools were used to form various analyses such as frequency lists, concordance lines and collocation patterns.

Voyant Tools is a web-based text reading and analysis environment. It is designed to make it easy for users to work with their texts in a variety of formats. Through Voyant, as shown in Figure 2, the students were able to examine the vocabulary density, readability index, average words per sentence, most frequent words in the corpus and their concordances.

Figure 2. Voyant corpus tool and sample analysis of the first episode of *My Name is Red* (I am a Corpse/Ben Ölüyüm) by Orhan Pamuk.

Lextutor is a corpus tool which is used widely for different purposes. For the purposes of this study, Text Concordance v.3.6 Eng. (https://www.lextutor.ca/conc/text/) was used. It is easy for inexperienced users and as is seen in Figure 3, it offers a list of words, phrases, or distributed structures and immediate contexts from a corpus or text (Web Concordancers, 2022).
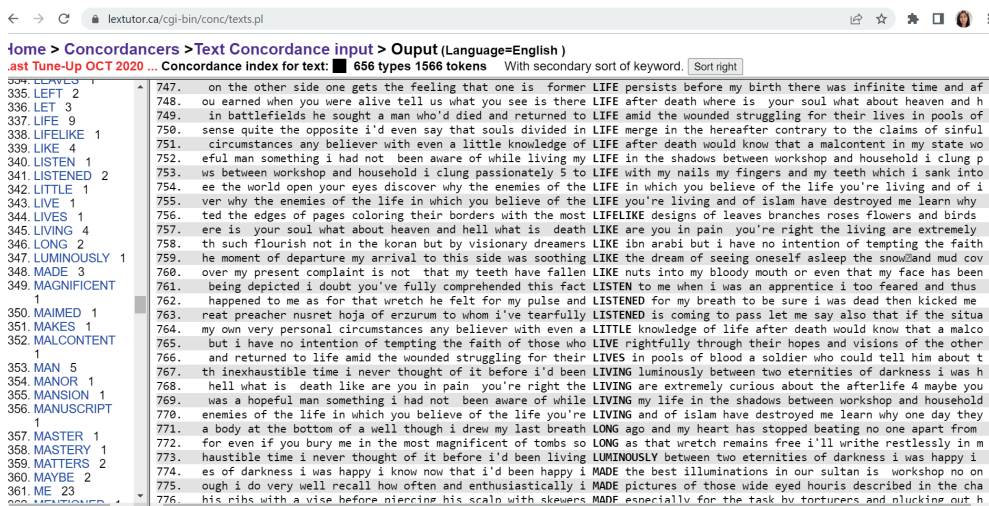
Figure 3. Lextutor corpus tool and sample analysis of the first episode of *My Name is Red* (I am a Corpse/Ben Ölüyüm) by Orhan Pamuk.

These tools have some advantages, for instance you do not need to align the texts before creating the corpus. The procedure is quite simple; it is necessary for users to copy and paste the text on the tools, and press the submit button to see the results of the analysis. Another advantage is that they are web-based tools, which means users do not need to download the tools on their computer, and they can copy or paste the text which will be planned to be the corpus and/or upload it in various text formats. The disadvantage is the necessity of internet connection.

For the assignments, firstly, the prologue and the first chapter of the book were chosen to be analysed in BA level, and six chapters (approximately 50 pages) of *Benim Adım Kırmızı* by Orhan Pamuk and its translation (My Name is Red) were assigned for MA level students. The corpus size of each assignment is displayed in Table 6.

Table 6. Corpus size of the assignments.

Level	Book	Words	Characters
BA level-Monolingual (source text) corpus	*The Suspicions of Mr Whicher or The Murder at the Road Hill House* by Kate Summerscale (Prologue and the first chapter of the book)	6158	34529
MA- Parallel Corpus	*Benim Adım Kırmızı* by Orhan Pamuk (Six chapters)	8120	61159
	My name is Red translated by Erdağ Göknar	12865	71855

Findings

The aim of this study is to explore how corpus tools can be used to analyse literary texts and their translations, and to implement methods of exploiting corpus tools in a literary translation context for the BA and MA level students of translation and interpreting in a university in Turkey. Additionally, the participants' views on using corpus tools were taken through an online survey.

Having concluded the 6-week procedures on stylistic analysis and using corpus tools, the participants were given assignments which cover stylistic analysis through Voyant and Lextutor in the BA and MA levels.

Thirty nine students' assignments were examined, and it is clearly seen that only two students did not use the corpus tools since they didn't participate in the lessons which included sample analysis together with the students. It is also clearly seen that participants use the tools only to analyse lexical and grammatical categories, which is an expected result due to the difficulty of finding figurative language features via corpus tools and/or they are not automatically extracted from the corpus. Through Voyant, it is possible to conduct searches according to various criteria, including lexical density, topic and grammar. Moreover, through Lextutor, it is possible to see concordancers and many more details.

Participants used the tools for different purposes. For instance, as Figure B1 displays the sample screen of Voyant from an assignment of a student (see Appendix B), Voyant was used for vocabulary density and frequent words, which serves for lexical category according to Leech and Short's (2007) stylistics analysis. The student, as all the students did, constructed a lexical analysis table using the information from Voyant (See Table B1 in Appendix B)). Figure B1 and B2 present the sample screen of Lextutor and Voyant which learners created (see Appendix B).

Moreover, participants were supposed to comment on each category and sub-category using all the answers to the questions in the checklist of the stylistics analysis. It is observed that most of the students (N. 25, 65%) failed to make a comment (meaningful comment) on the categories and sub-categories.

Furthermore, 3 MA level students analysed the style of *Benim Adım Kırmızı* by Orhan Pamuk and its translation of the first six chapters of the book using Lextutor and Voyant. After using the tools, they were assigned to further analyse the cohesive devices of the text, through Microsoft Excel, and they analysed the frequencies of *-dA (and, too, and what's more), gibi (like, as), ki (which, who,what, where) ve (and)* and their translations. Figure 6 shows the screenshot of the analysis –dA using Microsoft Excel. In this analysis, it is found in the classroom that, within the limitations of the small size corpus, Pamuk uses *-dA* as a connective more frequently than connective *ve* in arguments, which means he connects phrases with ve, while using -dA in arguments/sentences. Also, the translator mostly omitted them (-dA, ve) despite the fact that they play a significant role in cohesion of the source text. The connective

and in English and the Turkish equivalent *ve* are the most frequently used connectives, and they convey expansion relations, which are "used for relations that expand the discourse and move its narrative or exposition forward" as well as other relations such as temporal (PDTB 3, 2019, p.25). Therefore, they contribute highly to the cohesion of a text. Also, –dA is "generally described as a multi-function particle fulfilling the roles of focaliser, topicaliser, additive and intensifier, lexically carries an existential presupposition and a distinctness constraint on the value of at least one of the variables in the presupposition (Göksel and Özsoy, 2003).

The second research question was: *What are the views of students who use corpus tools to analyse literary texts and their translations?*

The survey of the study consists of 8 questions inquiring the level of helpfulness of corpus tools (see Appendix A) and 18 out of 39 participants responded to the short online questionnaire. They rated the helpfulness of the tools and exercises on the basis of the elements of style, respectively, in aiding preparation of literary analysis: lexical analysis (For example, verb, noun, adjective etc.); grammatical analysis (For example, sentence type, sentence complexity etc.); figures of speech (For example, metaphor, simile, repetition, personification etc.); cohesion and context analysis (For example, connectives, parallel forms etc.). Table 8 shows the level of helpfulness of the tools according to the categories of style. It is clear that participants find the tools *very helpful* when using lexical (N 9) and grammatical analysis (N 12), and they find it *helpful* when using lexical (N 9) and grammatical analysis (N 6).

Table 7. The helpfulness of the tools according to the categories of style.

	Very helpful	Helpful	Not helpful
Lexical analysis	9 (50%)	9(50%)	-
Grammatical analysis	12 (66.6%)	6 (33.3%)	-
Figures of speech	7 (38.8%)	7 (38.8%)	4 (22.2%)
Cohesion and context	2 (11.1%)	12 (66.6%)	4 (22.2%)

Eighteen participants answered the questionnaire and 67% (N 12) of them find the corpus tools *helpful* and 33% (N 6) find them *very helpful*, which means they all have positive views on the corpus tools used in this study (See Graph 1).
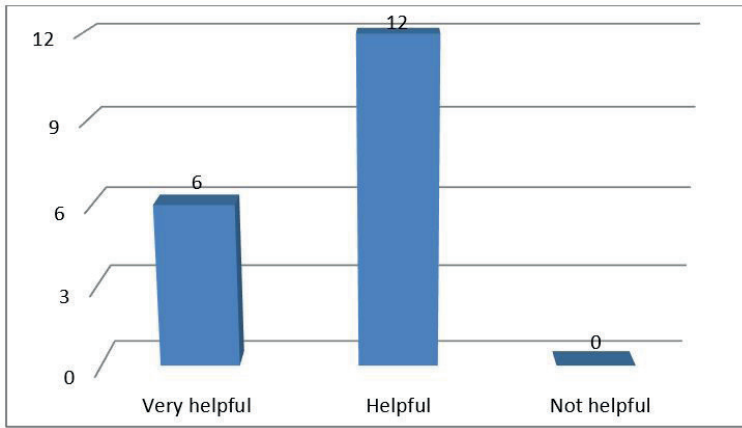
Graph 1. General helpfulness levels.

The next question was: *Do corpus tools you used help in your self-study, for instance, for other courses and/or for your translation projects?* While 13 participants responded it Yes, 5 participants responded No (see Graph 2).
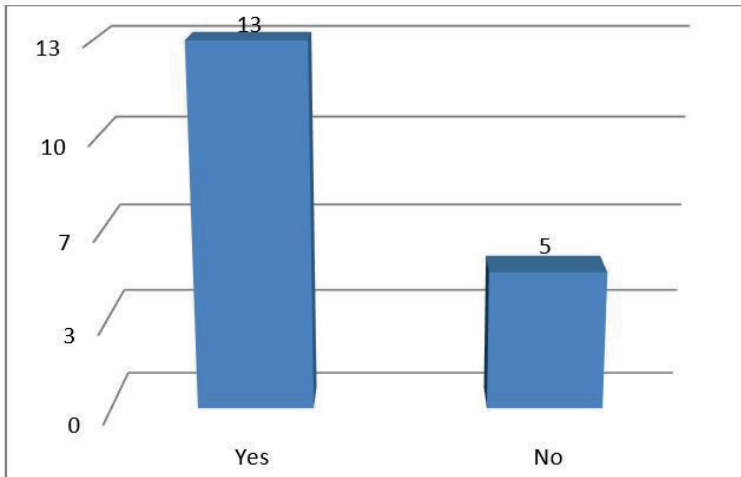
Graph 2. Do corpus tools you used help in your self-study?

The next question was: *What did you like and/or dislike about working with corpus-based tools?* Fourteen out of 18 participants wrote short comments. Table 9 displays the thematic analysis of the short comments of the participants on using corpus tools in stylistic analysis. Briefly, out of among 14 comments to the open ended item of the survey, participants used 25 positive words such as easy, useful, helpful etc., 9 negative words such as difficult, complex etc.

Table 8. Participants' short comments on using corpus tools in stylistic analysis.

Thematic analysis of short comments of the participants	
Positive	**Negative**
Convenient (1) *Easy (1)* *Easily (1)* *Easiest (1)* *Help (3)* *helpful (6)* *like (4)* *liked (2)* *phenomenal (1)* *quicker (1)* *useful (3)* *usefulness (1)*	*Complex (3)* *Complexity (2)* *Complicated (1)* *Difficult (1)* *Disliked (1)* *Hard (1)*
25 positive words	**9 negative words**

Conclusion

Returning to the questions posed at the beginning of this study, it is now possible to state that firstly, corpus tools can be used to analyse literary texts and their translations focusing on their styles. Secondly, the views of students who use corpus tools to analyse the style of literary texts and their translations are mostly positive, they generally find the process *very helpful* (N 6) and *helpful* (N 12).

The findings of this study have important implications for developing a method of stylistic analysis by using corpus tools in translator training.

An implication of this study is the possibility that corpus tools can be used for stylistic analysis in a clearly guided way in both BA and MA levels of translation and interpreting students, however, more effort is required to zoom into the analysis either manually or by using Microsoft Excel which may help to analyse the aligned parallel translation corpora. Although this can lead to a more complex process, it is necessary for a deeper and wider level analysis, as one was carried out on the MA level students.

The results of this study indicate that using corpus tools enhance the analysis of style of an author and/or a translator. Regarding figures of speech, although it is not possible to spot tropes, for instance metaphors automatically through corpus tools, more effort needs to be given to spot specific repetitions, forms of lexical and structural parallelisms (see, e.g. Şen Bartan, 2022), and repetitions of key words. It is possible to further guide the learners to analyse the repeated sequences of words (keywords) and interpret them. Moreover, it is possible to see manually which metaphors are used for instance, together with specific words, for instance, which metaphors are being used frequently together with "women/ men" as they represent gender biases. In this regard, it should be mentioned that students need to be aware of what to look for in terms of style.

Also, specifically BA level students have difficulty in making comments/interpretations on frequency analysis. This result may be explained by the fact that it needs more complex and sophisticated effort, and may be a "literature bank" to make general comments of the outcomes of the quantitative analysis through corpus tools. Therefore, not reading literary books, and not writing frequently could be a major factor, if not the only one, causing a lack of a general comment of the style of the author and/or the translator.

Furthermore, as the current study was limited to presenting the use of corpus tools for stylistic analysis in translator training, the effect of their literature knowledge was not estimated. Although scholars and instructors in the field of translation, language learning and literature need additional pedagogical tools such as corpus tools for stylistic analysis, students, also, always need to read as many novels, short stories and poems, or even nonfiction as possible in order to be exposed to literature. For this reason, a much more systematic study would identify how literature knowledge/experience interacts with the use of corpus tools in translator training.

REFERENCES/KAYNAKÇA

Akalın, R. (2022). Çeviri üzerine söylem üretmek ile çeviriyi uygulamak arasında: Çeviri alanında bilgi kazanımının seçenekleri üzerine bilimkuramsal bir sorgulama, *Ankara Üniversitesi Sosyal Bilimler Dergisi 13* (2), 183-192. http://dx.doi.org/10.33537/sobild.2022.13.2.16

Baker, M. 1993. Corpus linguistics and translation studies: Implications and applications. InM. Baker, G. Francis and E. Tognini-Bonelli (Eds.) *Text and technology: In honour of John Sinclair*, (pp.233-250), Amsterdam: John Benjamins.

Baker, M. (2000). Towards a methodology for investigating the style of a literary translator? *Target 12* (2): 241–266.

Bale, R. (2015). An evaluation of spoken corpus-based resources in undergraduate interpreting training. *International Journal of Applied Linguistics, 25* (1), 23-45.

Čermáková, A. (2015). Repetition in John Irving's novel A Widow for One Year: A corpus stylistics approach to literary translation. *International Journal of Corpus Linguistics, 20* (3), 355-77. https://doi.org/10.1075/ijcl.20.3.04cer

Corpas Pastor, G. (2018). Tools for interpreters: the challenges that lie ahead. *Current Trends in Translation Teaching and Learning E, 5*, 157 – 182.

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

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research.* CA: Sage. Göksel, A. & Özsoy, A.S. (2003). dA: a focus/topic associated clitic in Turkish. *Lingua 113.* 1143–1167.

Güner, S. P. (2018). *Derlem tabanlı yaklaşımların çeviribilmdeki yeri ve önemi*, [Basılmamış Doktora Tezi], İstanbul Üniversitesi. İstanbul.

Hansen, I., and M. Shlesinger (2007). The silver lining – technology and self-study in the interpreting classroom. *Interpreting 9* (1): 95–118.

Hu, K. (2016). *Introducing corpus-based translation studies.* New York: Springer Heidelberg.İnce, Ü. (2019). *Çeviri bilinci: Çevirenler, çeviremeyenler, çeviriverenler*. İstanbul: Tekin Yayınevi.

Kılıçkaya, T. & Ergil, B. (2023). The Importance of English Literature Courses in Translation Education and Its Place in Associate Degree Translation Programs in Turkey . *Söylem Filoloji Dergisi , Çeviribilim Özel Sayısı,* 243-259 . DOI: 10.29110/soylemdergi.1187079

Kurultay, T. (1995). Çevirmen, çevireceği metni nasıl okursa okumuş olur? Ya da çeviri sürecinde bir ön aşama olarak 'Çeviri amaçlı metin çözümlemesi' var mıdır? Işın Bengi Öner (Yay.), Çeviribilim 1. Çeviri Eğitiminde Özgün Metni Yorumlama, Çeviri Metni Oluşturma Sürecine Yönelik Yöntem Önerileri. Bursa: Ankara Üniversitesi TÖMER Bursa Şubesi, (s.111-124).

Laviosa, S. (2004). Corpus-based translation studies: Where does it come from? Where is it going? *TradTerm 10*: 29–57

Laviosa, S. (1998). The corpus-based approach: A new paradigm in translation studies. *Meta 43*(4): 474–479.

Leech, G., and Short, M. (2007). *Style in fiction: A linguistic introduction to English fictional prose* (Second Edition). London: Longman.

Mahlberg, M. (2013). Corpus analysis of literary texts. In Carole A. Chapelle (Ed.) *The Encyclopedia of Applied Linguistics*, Oxford: Wiley-Blackwell. (pp. 1282–1289)

Mahlberg, M., & Wiegand, V. (2020). Literary stylistics. In Adolphs, S., and Knight, D. (Eds.), *Routledge Handbook of English Language and Digital Humanities*. (pp. 306-327) Routledge.

Mastropierro, L. & Mahlberg, M. (2017). Key words and translated cohesion in Lovecrat's At the Mountains of Madness and one of its Italian translations. *English Text Construction 10*(1), 78–105. Amsterdam: John Benjamins.

McIntyre, D. (2013). Corpora and literature. In Carole A. Chapelle (Ed.) The Encyclopaedia of Applied Linguistics. Wiley-Blackwell.

Mikhailov, M. (2022). Text corpora, professional translators and translator training, *The Interpreter and Translator Trainer, 16*(2), 224-246, DOI: 10.1080/1750399X.2021.2001955

Nebot, E. M. (2003). Corpus-based Teaching: The Use of Original and Translated Texts in the Training of Legal Translators, *Translation Journal 7*(4), http://accurapid.com/journal/26edu.htm.

Nebot, E. M. (2008). Corpus-based activities in legal translator training, *The Interpreter and Translator Trainer (2)*2, 221-252, DOI: 10.1080/1750399X.2008.10798775 https://doi.org/10.1080/1750399X.2008.10798775.

Nord, C. (1991). *Text Analysis in Translation: Theory, Methodology and Didactic Application ofa Model for Translation-Oriented Analysis*. Amsterdam: Rodopi.

Olohan, M. (2003). How frequent are the contractions? A study of contracted forms in the Translational English Corpus. *Target 15*(1): 59–89.

Olohan, M. (2004). *Introducing corpora in translation studies*. London: Routledge.

Reppen, R. (2010). *Using corpora in the language classroom* (Cambridge language education). New York: Cambridge University Press.

Pearson, J. (2003) Using parallel texts in the translator training environment. In F. Zanettin, S. Bernardini, and D. Stewart (eds.), Corpora in translator education. Manchester: St. Jerome, 15–24.

Sevimay, F. (2018). *Çeviri'Bilirsin! Edebiyatın gizli kahramanlığı hakkında notlar.* İstanbul: Hep Kitap.

Şen Bartan, Ö. (2022). Structural repetitions and discourse relations in English-Turkish translations of TED talks. *The Literacy Trek,8*(2), 186-202. https://doi.org/10.47216/literacytrek.1119118.

Tutaş, N. (2006). Yazınsal metinlerde biçembilimsel inceleme, *Littera Edebiyat Yazıları 19*, ss.169-180.

Voyant Tool Help (2022). https://voyant-tools.org/docs/#!/guide/start

Webber B., Prasad Rashmi, Lee Alan, Joshi Aravind. (2019). The Penn Discourse Treebank 3.0 Annotation Manual. Philadelphia, University of Pennsylvania, https://www.seas.upenn.edu/~pdtb/.

Wolcott, H. F. (1994). *Transforming qualitative data.* Thousand Oaks, CA: Sage.

Zanettin, F. (1994). Parallel words: Designing a bilingual database for translation activities. In A. Wilson & T. McEnery, (Eds.), Corpora in language education and research: A selection of papers from Talc94 (Technical Papers, Volume 4; pp. 99-111). Lancaster, UK: UCREL.

Zanettin, F. (2007). Parallel Corpora in Translation Studies. Issues in Corpus Design and Analysis, In Wolfgang Teubert and Ramesh Krishnamurthy (Eds.) Corpus Linguistics: Critical Concepts in Linguistics, London & New York: Routledge, 285-98.

Appendix A
Questionnaire

The questionnaire will take approximately 5 minutes.

1. What is your native language?

2. Please rate the helpfulness of lexical analysis (e.g., verb, noun, adjective, etc.) of the author's/translator's style through corpus tools such as Lextutor, Voyant, etc. in aiding your preparation of literary translation.

   Very helpful

   Helpful

   Not helpful

3. Please rate the helpfulness of grammatical analysis (e.g., sentence type, sentence complexity, etc.) of the author's/translator's style through corpus tools such as Lextutor, Voyant, etc. in aiding your preparation of literary translation.

4. Please rate the helpfulness of figures of speech (e.g., Metaphor, simile, repetition, personification etc.) analysis of the author's/translator's style through corpus tools such as Lextutor, Voyant, etc. in aiding your preparation of literary translation.

5. Please rate the helpfulness of cohesion analysis (e.g., connectives, parallel forms etc.) of the author's/translator's style through corpus tools such as Lextutor, Voyant, etc. in aiding your preparation of literary translation.

6. Please rate the helpfulness of generally stylistic analysis of the author's/translator's style through corpus tools such as Lextutor, Voyant, etc. in aiding your preparation of literary translation.

7. Do corpus tools (such as Lextutor and Voyant) you used help in your self-study, for instance, for other courses and/or for your translation projects.

8. What did you like and dislike about working with corpus-based materials/tools (such as Lextutor, Voyant)?

Appendix B
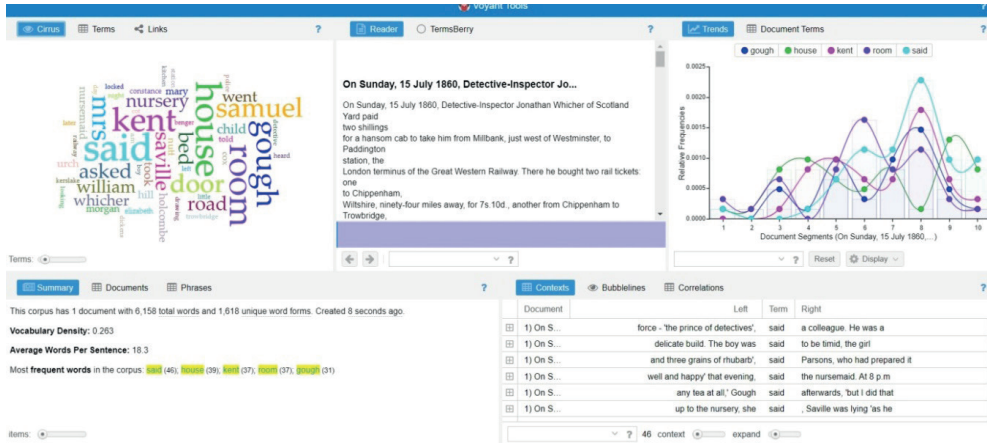## Tables and figures presenting students' sample analysis

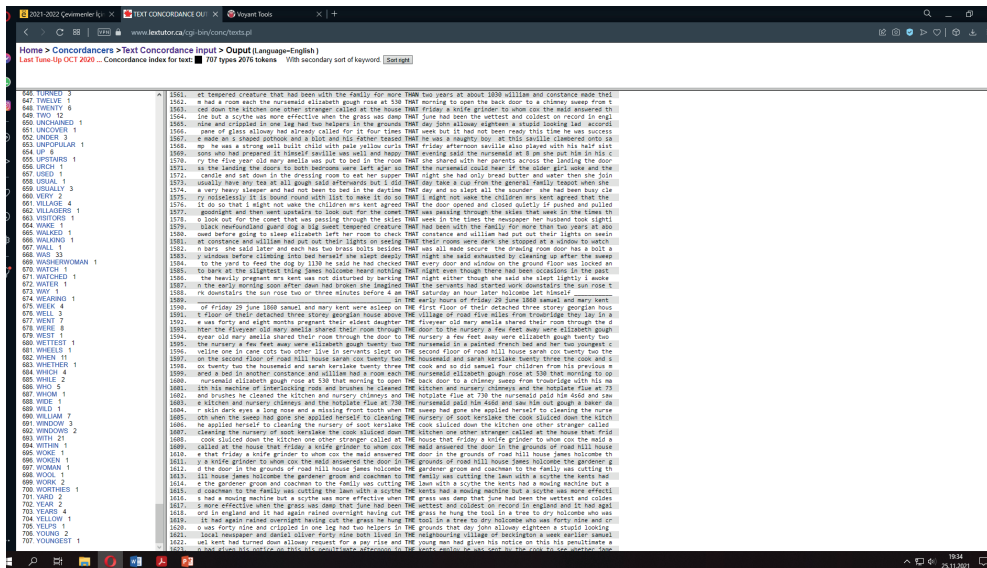**Figure B1.** A learner's screenshot of Voyant output.

Figure B2. Learner's screenshot of Lextutor output.

Table B1. The analysis of lexical categories by a BA student.

Nouns	Adjectives	Adverbs	Verbs
Bed (21)	Alone	Ordinarily	Was (120)
Arms (3)	Bad (2)	Usually (3)	Said (68)
Blood (12)	Big	Very (5)	Asked (23)
Body (10)	Cold	Noiselessly	Answered (4)
Boots (5)	Confused (2)	Quietly	Believe
Boy (18)	Dark (3)	Sweetly	Lay (3)
Candle (5)	Delicate	Mainly	Find (3)
Child (20)	Difficult	Deeply	Rose (4)
Children (11)	Elder	Gently	Cleaned (3)
Corpse (5)	Empty (2)	Shortly	Search (9)
Cot (10)	Fast	Probably (2)	Saw (5)
Day (8)	Good (2)	Heavily	Doodled
Door (28)	Happy	Lightly	Drive (2)
Room (39)	Heavy (2)	Frequently	Played
Window (9)	Large	Certainly (2)	Pay (2)
Village (6)	Late	Directly	Left (13)
Traces (4)	Little (13)	Harshly	Carried (5)
Throat (3)	Long (2)	Immediately	Put (11)
Blanket (16)	Peaceful	Smoothly	Took (15)
Son (8)	Reluctant	Particularly	Looked (8)
House (41)	Sharp	Lively	Pulled (4)

Table B2. General Comments on lexical categories made by a BA student (spelling mistakes and other errors were kept).

General: The sentences are written in colloquial language. The sentences are little bit complex due to Relative clause structures and reported speech. The sentences are descriptive and a little bit specific due to the terms which are used such as "PC Urch, membranes, blood vessels, nerve vessels and air tubes." (Summrscale,2008)

Nouns: Nouns are mainly concrete and they refer events.

Adjectives: Adjectives are frequent. They are descriptive, but there are some of which refer degree such as "bad" and "good". They are also attributive.

Adverbs: Adverbs are also frequent but they are not frequent as much as adjectives. There are adverbs which refer time, such as "Generally"; there are also adverbs which refer manner (they are mainly refer manner), such as "lively, smoothly, harshly..." and so on.

Verbs: Verbs are dynamic. They carry an important meaning in the story and they generally refer movements. They also refer speech acts, such as "answered, said, asked, replied."