Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Triennale Interclasse in
Lingue, Letterature e Mediazione Culturale (LTLLM)
Classe LT-11

Tesina di Laurea

Misogyny in Oscar Wilde's The Picture of Dorian Gray*: A Corpus Stylistic Analysis*

Relatrice
Prof. Katherine Ackerley

Laureanda
Alice Costantin
n° matr.1199818 / LTLLM

Anno Accademico 2021 / 2022

Table of contents

Introduction Page 4

1. Corpus stylistics 6

1.1. Corpus stylistics: a combination of stylistics and corpus linguistics 6

1.2. Approaches and methodologies 10

1.3. How different disciplines analyse literary texts and the benefits of corpus stylistic studies 14

2. Oscar Wilde, *The Picture of Dorian Gray*, and other researchers' corpus stylistic analyses of his works 18

2.1. Oscar Wilde's biography and his novel *The Picture of Dorian Gray* 18

2.2. Misogyny and representation of women in *The Picture of Dorian Gray* 21

2.3. Overview of corpus stylistic analyses of Wilde's works 25

3. My corpus stylistic analysis of Wilde's *The Picture of Dorian Gray* 31

3.1. Introduction of my analysis: corpora, tools, and expectations 31

3.1.1. Corpora 31

3.1.2. Methodology 33

3.2. Results 34

3.3. Comparing and contrasting expectations with findings 51

Conclusion 54

Bibliography 56

Summary in Italian 59

Introduction

The aim of my thesis is to observe and analyse Oscar Wilde's novel *The Picture of Dorian Gray* (1891) and its misogynistic aspects with the aid of corpus techniques. The choice of this literary work was determined by the fact that this is one my favourite books of all time, and ever since I read it, I have been impressed by some of the words used by the author against women. Thus, after becoming familiar with *AntConc*, a concordance program, during the first year of my Bachelor's Degree, I was curious to investigate this aspect of the novel through a corpus stylistic study. More specifically, I will use a corpus-assisted approach to examine Wilde's language and literary choices referring to women, when and how his words express or denote, for example, prejudices and insults against women. Therefore, I will observe forms of sexism that keep women at a lower social status than men, which in other words means maintaining the societal role of patriarchy. In the present study, two corpora are used: the target corpus is made up by Wilde's *The Picture of Dorian Gray*, and the reference corpus, *Wilde*, comprehends 15 texts (collections of essays, collections of fairy tales, letters, plays, poetry, short stories, and travel descriptions) that were all written by the homonymous author. The decision and selection of the works to include in the reference corpus were influenced by the fact that they all share the same author, and I wanted to observe how he dealt with the same topic in different productions of his, and I also wanted to investigate whether his only novel, *The Picture of Dorian Gray*, could be one of his most misogynistic works.

The first chapter of the present thesis is the most theoretical one, since I clarify what corpus stylistics is, providing a brief theoretical overview of this emerging discipline, and in doing so, I also explain what two other disciplines, corpus linguistics and stylistics, since their combination is precisely what lead to the formation of corpus stylistics. In addition, I describe its main features and goals, as well as the variety of methodologies and approaches that can be used conducting a corpus stylistic analysis. In this chapter I even explain why corpus stylistic analyses can be beneficial, comparing and contrasting their results with the ones produced by linguistic or literary studies, which are other disciplines that investigate literary texts.

In the second chapter of my work, I provide a brief biography of Oscar Wilde, and I summarise the plot of his only novel, *The Picture of Dorian Gray*. In this chapter I also report and quote several examples of misogynistic remarks, sexist comments, and stereotypes about women that can be found in the novel. In order to so, I thought that it was fundamental to acknowledge the influence of Wilde's own society, and thus I illustrate the social background of the Victorian period, highlighting its typical sexist attitude, which is clearly reflected on the novel. Finally, in this chapter I dedicate a section to report and discuss other researcher's corpus stylistic analyses that are similar to the one that I want to conduct.

The third and final chapter of my thesis entails the corpus stylistic analysis. After introducing the two corpora and the tools that I chose to use in my work, and after explaining what *AntConc* is and how it works, I state what I expect to find as regards the results of my research question. Indeed, since I intend to conduct a corpus-assisted analysis, which tests preconceived hypotheses, I presume that women are mentioned several times in the target corpus. Moreover, I expect to find many adjectives that describe physical appearance, beauty, and charm. This is due to the fact that when I read the novel, I had the impression that many conversations are about these topics, especially when male characters talk about women, maybe even in a bad way. I also believe that women are not appreciated or celebrated for their intelligence. Hence, more generally, I think that negative contexts are often present in the novel, mostly in passages in which women are criticized or insulted. In the following step of my thesis, I present the results and I reflect on them, comparing and contrasting them with my initial expectations.

Chapter 1

Corpus stylistics

The present chapter provides a brief theoretical overview of corpus stylistics, its main features, its goals, and its most used methods and approaches. I begin by introducing what corpus stylistics is, what this specific field of study consists in, and what characterizes it, also describing the disciplines of corpus linguistics and stylistics. Subsequently, I focus on these methods employed in the study of literary texts and how they work, and I offer some examples of the possible approaches to the analysis of literary works with the aid of corpus techniques. The chapter concludes by outlining how different subjects, such as corpus stylistics, linguistics, and literary studies, can be applied to analyse literary texts in dissimilar ways. In this final part I also highlight the strengths and the innovations resulting from the use of software in the analysis.

1.1. Corpus stylistics: a combination of stylistics and corpus linguistics

In this section I explain what corpus stylistics is, and in order to do so I also present two other important fields of study: stylistics and corpus linguistics. More specifically, this initial theoretical part clarifies how corpus stylistics works, what it does, and what the key characteristics and aims of this discipline are.

Corpus stylistics is a relatively new field of study that combines two disciplines: stylistics and corpus linguistics. Indeed, the recent studies carried out under the umbrella of 'corpus stylistics' integrate the methods of corpus linguistics and the goals and methods of traditional stylistic research.

Stylistics is the linguistic analysis of literary texts, and its focus is to linguistically analyse a literary work to decode literary meanings as well as structural features by separating linguistic patterns and their functions in a text. In this linguistic discipline, which is the only one to allow the analysis of literary texts and their meanings by way of linguistic techniques, the connection between linguistic form and meaning, mostly of literary texts,

is a crucial point. In addition, unlike other linguistic disciplines that usually employ only one set of analytic tools, stylistics is not a closed discipline with a limited number of analytic techniques at its disposal, since this subject uses all analytic techniques of linguistic text analysis, and this is one of its strengths.

On the other hand, corpus linguistics is a field of linguistics engaged in the electronic analysis of language data, and it assumes a correlation between the frequency of a pattern and its significance in the data (Fischer-Starcke, 2010: 3). Corpus linguistic research is based on the analysis of a corpus, which is a large and principled collection of texts stored on computer (Biber, 2011: 15). In other terms, a corpus is a large collection of texts designed for a specific purpose while also representing a particular language or sub-language. Insights into the language system are possible only through the analysis of more than a single text, and it is exactly the analysis of large amounts of language data that makes studying actual language use possible. Research carried out on a corpus aims at describing the language patterns used in the target text: corpus linguistics gains insights into the language of a specific text or corpus and into the language variety represented by that corpus. This is possible because these studies, as explained by Cvrček and Fidler (2019), demand the use of two corpora: a target corpus, which includes the set of texts to be analysed, and a wider, more general comparison corpus, which is also known as reference corpus. The former includes the literary work(s) in question, and the latter is required to represent 'typical' patterns of use, making it possible to empirically recognize distinctive linguistic patterns in the target corpus.

Fischer-Starcke (2010: 19) identified four criteria for the validity of corpus linguistic analyses applied to stylistics, which are designed to allow the evaluation of the applicability and use of corpus linguistic techniques in the analysis of literary works. The criteria are:

1. Growth of knowledge

   Research is noteworthy either when it proves the usefulness of employing corpus linguistic techniques to stylistics or when their adoption is rejected. In other words, the techniques are successful when literary insights or additional information on the data are gained precisely because of computerised analyses, or

when already known information on the data or its previous interpretations can be supported or refuted thanks to the electronically generated data.

2. Replicability of results

   Researchers can potentially repeat what others have done to test their own results. This possibility makes the method, the technique, and the result of the research transparent, and transparency grants the replicability of an analysis, also revealing its subjective elements, which cannot be avoided. The fact that an analysis is transparent enables other researchers to question what has been previously done and gives the opportunity of deciding differently in their own work. Hence, it allows the evaluation of an analysis as a whole.

3. Checkability of results

   This criterion is linked to the previous one: analysis results must be coherent, and it is crucial that they can also be tested by other experts to assert their validity. But, as stated above, checking other analysts' works and conclusions is possible only if the analyses are transparent. Thus, when transparency is provided, results obtained thanks to corpus linguistic analyses can be corroborated or doubted, but neither verified nor falsified, due to the probabilistic nature of this discipline. To put it differently, the outcome of the analyses describes solely tendencies and probabilities in the language, and they are the consequences of generalizations of data extracted from a text or corpus by way of electronic means. Nevertheless, generalizations regarding a context can only be extensions of the data type in the original analysis, meaning that statements on language exclusively apply to the language variety present in the considered corpus, but not to general language usage. Indeed, language is an open system, and a corpus usually represents only an excerpt of the possible language of the analysed variety. Consequently, generalizations on language mostly apply to features that occur in the corpus, they show the probability with which that feature occurs in the language variety represented by that particular corpus. Generalizations on linguistic patterns from a representative corpus are one of the potentials of and, indeed, an explicit goal in corpus linguistics (Fischer-Starcke, 2010: 14).

   However, the difference between replicability and checkability lies in the former referring to the reproduction of analyses, unlike checkability, which refers to

testing the method and techniques by conducting one's own research using some of the parameters of the original study. Therefore, replicability ensures that the techniques used, and the decisions taken are transparent to other researchers. Checkability, on the other hand, emphasizes the transfer and adaptation of the techniques to different sets of data, different analytic techniques, or different research questions (Fischer-Starcke, 2010: 23).

4. Innovation derived from analyses

   This criterion questions whether analytic techniques are innovative and whether they contribute and lead to a growth of knowledge in corpus linguistics, stylistics, or literary studies. However, not everything that furthers knowledge is necessarily new and innovative, given that non-innovative conclusions might still be useful for specific questions.

The four criteria for evaluating research can be fulfilled uniquely by corpus linguistics and corpus stylistics, but only partly by literary studies. This is due especially to replicability and checkability, which can be applied exclusively to systematic analyses, since they are the ones that can be compared to other analyses, and consequently evaluated in the context of other studies.

In corpus stylistics, conclusions about the meanings of the data are based on the assumption that form and meaning are correlated. Not by chance, the assumption of correlation between the frequency of a linguistic feature and its significance in a text or corpus fundamental. However, the frequency of a pattern or feature varies between different corpora and texts. Whether a certain linguistic pattern occurs in the data depends on several factors, such as sociolinguistic elements, differences between written and spoken language, genre conventions, and/or linguistics.

In general, frequent features are considered more relevant than rare ones in the analyses. Frequency is an indicator for typicality of language usage, that is, frequent linguistic features are relevant in terms of style of writing of a particular text or author, and style is the typical language of a text.

Therefore, the term ‘corpus stylistics’ refers to the use of corpora, and the tools and methods of corpus linguistics, applied to the study of literature. The fact that corpora are stored electronically is exactly what makes it possible to process the texts with the help of computer tools. Other major goals pursued by the subbranch of corpus stylistics are the observation of how meaning is encoded in language and the development of legitimate working techniques to decipher these meanings, and the study of literary meanings of texts.

1.2. Approaches and methodologies

This subchapter is dedicated to the illustration of the variety of approaches and methods that researchers can use when conducting a corpus stylistic study. In addition, I give some examples of what some analysts have found through these procedures that mainly examine repetitions and word distributions in certain literary texts. These discoveries and observations are relevant to my study because I expect to conduct a similar analysis in my own target text.

There are different methodological techniques that can be applied to corpus investigations, but corpus analysis usually shares four characteristics:

1. It is empirical, meaning that it analyses the actual patterns of language.
2. It employs a corpus.
3. It makes extensive use of computers for the analysis, resorting to both automatic and interactive techniques.
4. It depends on quantitative techniques as well as qualitative analytical ones.

Over the last few decades, there has been an increasing number of stylistic studies of literary language involving computational analyses of words in literary texts. These works can be identified as ‘corpus stylistic studies’. These analyses have been using two major analytical approaches, which are commonly referred to as corpus-assisted and corpus-driven.

A corpus-assisted approach, in Tognini-Bonelli’s words (2001: 65), “avails itself of the corpus mainly to expound, test or exemplify theories and descriptions that were formulated before large corpora became available to inform language study”. This means

that this methodology uses corpus data to explore preconceived theories and hypotheses with the aim of validating, refuting, or refining them. This factor highlights that a corpus-assisted approach seems to have a limited potential in terms of innovative findings, as certain patterns might have not been detected before the introduction of corpus methods. In fact, this kind of analysis may be carried out in order to assess a stylistician's intuition about the stylistic effects of a specific text.

On the other hand, a corpus-driven approach is more inductive. In other words, the corpus itself is the basis, the source of hypotheses about language, and, as Biber states (2012: 4), "the linguistic constructs themselves emerge from analysis of a corpus". This second approach reduces a priori assumptions as regards linguistic features, attempting to uncover new linguistic constructs through a more inductive analysis of corpora. Thus, in a corpus-driven analysis, the analyst must 'go in blind', that is, without presuppositions about the data.

Most corpus stylistic studies focus on the distribution of words to identify textual features that are characteristic of an author, a specific text, or even of a literary character. Corpus software is used to quantify linguistic phenomena and display data so that the researcher can investigate linguistic patterns (Mahlberg, 2010: 292). Indeed, corpus software collects and accurately organises large amounts of textual data. Evans (2018: 2) explains that corpus software "reorders the items in a corpus so that they can be observed and investigated by the user and it calculates statistical information about the data in the corpus". *AntConc*, *kfNgram*, *Wmatrix*, *WordSmith Tools* are some examples of corpora software used in many corpus studies.

There are three main methodological approaches that have been used to explore and examine the stylistic relevance of word distributions: keyword analysis, the identification of typical extended lexical phrases, and collocational analysis.

Generating keywords can be a great starting point for the analysis of a novel, since this step provides a first overview of a text, highlighting the distinctive words that may be useful for a more in-depth analysis of a specific literary work. For example, Scott and Tribble (2006) did a keyword analysis of the characteristic words used in Shakespeare's *Romeo and Juliet*. In this case the researchers resorted to a reference corpus that consisted

of all the other plays written by Shakespeare. Culpeper (2002) even furthered this kind of analysis by examining the distinctive keywords of each character in *Romeo and Juliet*. In order to do this, Culpeper used reference corpora that consisted of the utterances pronounced by all the characters. This approach resulted in both expected findings (e.g., Romeo uses the words *beauty* and *love* more than other characters) and surprising findings. Thus, for Juliet, the analysis shows that the most distinctive word (i.e., the word with the highest 'keyness' score) is the conditional subordinator *if* (Biber, 2011: 17). This notion does not add new information as regards Juliet's mental state, but it does help to understand the linguistic devices and choices used to convey it.

Many corpus stylistic analyses focus on extended lexical phrases that are characteristic of authors or literary works, referred to as lexical 'clusters' or 'bundles'. A key-cluster analysis can be relevant to recognize which lexical sequences are prevalent in a specific text by comparing it to a corpus that is more general. The parameters for the length and minimum frequency of a cluster vary according to the research question (Mahlberg, 2010: 297). For instance, Mahlberg (2007) analysed clusters for a study of texts written by Charles Dickens, focusing on five-word clusters, and she noticed that *as if he had been* and *his hands in his pockets* are frequent lexical sequences of words in Dickens's novels, but they are not as common in a more general corpus of 19th century fiction. One specific group of clusters contains what Mahlberg (2010) calls 'labels', which are clusters containing names or expressions used to refer to characters, or clusters that are specific to individual texts. Mahlberg (2007) found labels occurring in Dickens's texts, noticing that *the father of the marshalsea*, *the person of the house*, and *the lady of the caravan* are the most recurrent ones. Thus, by observing these examples, it is also possible to see that labels overlap to some extent with what has been named in literature as habitual phrases that aid characterization, since this group of clusters is strictly linked to specific characters and concepts.

Yet another major approach used for corpus stylistic analyses concerns the notion of collocation, which is the way in which certain words are associated with each other, and they tend to co-occur more frequently than what would be expected by chance. Some cases have shown that a certain word can co-occur with collocates that have the same 'semantic prosody'. Semantic prosody refers to the attitudinal meaning, often pragmatic, of a lexical item (Cheng, 2013: 1). In other terms, a word might appear as neutral, but it

actually can be perceived with positive or negative meanings. Sinclair (1987) shows that the verb *set in* has a negative prosody, since it is usually associated with unpleasant events. Indeed, some of the nouns that co-occur with this verb are *rot*, *decay*, *ill-will*, and others.

In addition, the study of body language is another interesting aspect that can be observed in these analyses. One of the approaches suggested by Mahlberg and Smith (2010) is looking for 'suspensions'. A suspension is a span of (narrator) text which interrupts a span of quoted speech, thought, or writing (Mahlberg, 2010: 296). Suspensions are useful due to the fact that they contain information accompanying speech, and when people speak body language is always involved, it is unavoidable. Even though the extent to which body language is made explicit can vary in novels, contexts of speech provide helpful indications to find examples. A suspension identified by Mahlberg and Smith (2010) is highlighted in italics: '"And this," *cried Darcy, as he walked with quick steps across the room,* "is your opinion of me! …"'.

As can be noticed, the corpus stylistic approaches discussed above focus on words. In these studies, it is less common to analyse grammatical patterns, and there are two major reasons for this. First, analysing grammatical features in a corpus is a difficult process, since it requires specialised computer programs, such as a grammatical 'tagger'. But additionally, it is easier to notice the stylistic importance of word choice, while grammatical characteristics are much less salient (Biber, 2011: 19).

The approaches and methodologies that I presented in this section are only some of the possible procedures to the study of literary texts with the aid of corpus methods. These examples show that the application of corpus techniques to the study of literary texts requires the combination of quantitative and qualitative analyses in order to provide useful and complete insights. Indeed, literary scholars might take advantage of corpus methods, whereas corpus linguists might develop more specific corpus resources and methods due to the engagement with literary criticism.

1.3. How different disciplines analyse literary texts and the benefits of corpus stylistic studies

The last segment of this theoretical chapter focuses on how the disciplines of corpus stylistics, linguistics, and literary studies approach the analysis of literary texts. I do this especially by highlighting that corpus stylistic analyses can be more beneficial, complete, and efficient than the investigations covered by either linguistics or literary studies.

Linguistics and literary studies are both subjects that observe and analyse texts and their meanings. However, these two disciplines differ in their methods of analysis as well as in their choice of texts. In literary studies the literary value of a text is what often functions as a criterion for it to be selected for an analysis. As one may expect, literary studies are restricted to literary texts and the analysis of their meanings, unlike linguistic studies, which usually analyse non-fiction texts, fragments and collections of texts and the functions of patterns in the language of data. Moreover, literary studies gain knowledge mainly of a particular text, literary period, or author, while linguistics gains knowledge of the specific text or corpus, and of the language system with its mechanisms for encoding meaning. Yet another difference between these two subjects lies in their definitions of style: literary studies perceive style as an aesthetic choice, whereas linguistics perceives it as a choice of the author, since, as mentioned by Fischer-Starcke (2010: 5), "a text represents the results of a complicated selection process, and each selection has meaning by virtue of all other selections which might have been made, but have been rejected (Sinclair, 1965: 76f)".

But corpus stylistic insights have been included in interpretative textual analysis, and as previously mentioned, this text-internal analysis gives a new perspective on the data. Indeed, the detailed and systematic linguistic analysis can lead to the discovery of new meanings, which, as Fischer-Starcke affirms (2010: 6), are "virtually invisible in an intuitive approach to the data as literary studies". Furthermore, since corpus stylistics allows the analysis of complete texts or corpora, it provides the means to analyse larger sets of data using corpus linguistic techniques. A computer-aided analysis can help the reader to notice and keep track of the multiple patterns of the narrative, maybe even

alerting the analyst to subtleties and complexities that may otherwise have gone unnoticed. As a matter of fact, corpus linguistic techniques permit a systematic, detailed analysis of large quantities of language data for lexical and/or grammatical patterns, and they also allow to subsequently decode the meanings of these patterns, which, once again, are not intuitively recognizable due to the large size of the data.

Thus, new understanding of the data can be gained because the data is examined in a detailed and systematic way, but also because a larger number of units of meaning are taken into account. This happens because literary studies look at a text as a unit of meaning, while on the contrary corpus linguistic studies look at more than a single unit as carriers of meaning. Words, phrases, text parts, and the text itself are units of meaning which all contribute to the literary meanings of the data (Fischer-Starcke, 2010: 11-12). Of course, analysing more than one of these units of meaning leads to a more thorough view of the data. It is important to understand that this multi-dimensional approach is possible only because software is involved in the analyses, and this explains why multi-dimensionality is achievable exclusively in corpus linguistic and corpus stylistic works.

Moreover, Fischer-Starcke (2010: 16) highlights that the electronic analysis of data in both corpus linguistics and corpus stylistics seems to fulfil the goal of science and research, which is objectivity. This is the reason why this field of study is often referred to as the 'scientific study of literature'. A depersonalized software explores the data in a way in which the output of the software seems to have been generated without human interference. As a matter of fact, at first glance, corpus studies seem to be objective, but when looking at these analyses more closely, subjective elements do emerge. For instance, the choice of data and software for the analysis, the settings of the software, the choice of which of the data generated by the software is to be considered for the study, and, of course, the interpretation of the data.

Therefore, in corpus analyses both subjective and objective elements are present. In this discipline, subjectivity and objectivity are interdependent elements, they influence each other, and they cannot be separated. This results in the impossibility to achieve entirely objective results, due to the fact that every research project includes subjective features, especially the analyst's choices and interpretation of the data, but even other characteristics. At this point, in order to evaluate the meaning and relevance of every

single analysis, identifying and making explicit its subjective and objective components is not only helpful, but also pivotal, if dealing with corpus analyses.

In conclusion, since corpus stylistics, as Mahlberg states (2010: 295), "employs methods and approaches of corpus linguistics and links them with concerns in literary stylistics and literary criticism", in this field the techniques of corpus linguistics and the goals of stylistics complement each other as they unravel linguistic patterns and the meaning they have in a particular text. The two disciplines use the knowledge gained to generate a more general understanding of, for instance, literary meanings or the patterns of language. Thus, it is precisely the combination of corpus linguistics and stylistics that grants the decoding of the meanings of literary texts that cannot be detected either by intuitive techniques, as in literary studies, or with the necessary restriction to short texts or extracts, as in traditional stylistics.

Although the use of corpus techniques in stylistics has increased only in recent years, this practice has already demonstrated one of its major benefits. Unlike traditional stylistics, corpus stylistics enables the analysis of prose texts, making it possible to examine whole texts instead of only small extracts. This means that corpus stylistics allows the analysis of complete texts or corpora for their lexical, phraseological, and grammatical patterns. The ability to look at and analyse full-length texts facilitates the interpretations of literary works by describing an author's or a text's most representative features. Indeed, author's idiolects have always been one of the main concerns in corpus stylistics. These results can be achieved by identifying and interpreting linguistic patterns in a text or corpus. This procedure ensures that statements regarding the contents or structure of a text or corpus are based on the actual language of the data. The fact that many stylisticians see an advantage in corpus stylistics as regards the analysis of texts in their entirety highlights another important aspect, which is that this emerging discipline adds innovation to both stylistics and corpus linguistics.

This chapter is relevant to my study as it introduces the basic theoretical points relevant to the purpose of my research, since I intend to analyse Oscar Wilde's *The Picture of Dorian Gray* using corpus stylistic methods. In the following chapter, I will provide

information regarding the novel and its author, as well as misogyny present in this literary work.

Chapter 2

Oscar Wilde, *The Picture of Dorian Gray*, and other researchers' corpus stylistic analyses of his works

The second chapter of this work presents Oscar Wilde, since I intend to examine his masterpiece, *The Picture of Dorian Gray*, employing corpus stylistic methods. After providing a brief biography and summarising the plot of the novel, I continue by reporting some of the several examples of misogyny, female stereotyping, and sexism present in the considered novel. In doing so, I highlight which characters usually deliver misogynistic comments and hate speeches, and I also acknowledge the social background of the Victorian period, which clearly has an influence on the style of the novel. I conclude this chapter with some examples of corpus stylistic analyses conducted on similar topics or on other literary works of Oscar Wilde's, which I expect to be helpful for my research.

2.1. Oscar Wilde's biography and his novel *The Picture of Dorian Gray*

The first section of this chapter focuses on Oscar Wilde and his first and only novel, *The Picture of Dorian Gray*. I begin with a brief biography of the author, highlighting the most important moments of his literary and social career. In this subchapter I also mention some of his most successful works, especially Wilde's novel, which I summarise in a nutshell.

Oscar Wilde (1854-1900) was an Irish novelist, dramatist, poet, critic, and a celebrity, indeed, he was a spokesman for the late 19th century Aesthetic movement in England. His reputation mainly rests on his only novel, *The Picture of Dorian Gray* (1890), and on his comic masterpieces *Lady Windermere's Fan* (1892) and *The Importance of being Earnest* (1895). Besides literary accomplishments, he is also famous, or perhaps infamous, for his wit, flamboyance, and affairs with men (Tobin, 2019).

Born in Dublin in 1854, Wilde was educated at Trinity College, Dublin (1871-74), and Magdalen College, Oxford (1874-78). Oxford was crucial in Wilde's artistic and

philosophical development, since there he became involved in the Aesthetic Movement, being, as stated by Beckson (2021: 1), "deeply impressed by the teachings of the English writers John Ruskin and Walter Pater on the central importance of art in life and particularly by the latter's stress on the aesthetic intensity by which life should be lived". After graduating, Wilde moved to London to pursue a literary career, but in the early 1880s, he established himself in social and artistic circles, especially due to his wit and flamboyant personality and lifestyle, quickly winning a reputation as a brilliant conversationalist, dandy, aesthete, and fashionable society figure. Grinstein (1980) remarks that these qualities "even attracted the attention of writers and cartoonists", since with their operetta *Patience*, Gilbert and Sullivan made Wilde even more popular by basing the character of Bunthorne, a "fleshly poet", partly on him. The Irish playwright was not put off by these parodies, in fact, he appreciated that they provided him notoriety. Moreover, in order to reinforce this association, Wilde published a collection of poetry, *Poems* (1881), which received mixed reviews by critics. He was also invited to the United States and Canada for a series of lectures to exhort Americans to love beauty and art. When he returned to England, Wilde wrote a series of fables, which are collected in two volumes: *The Happy Prince and Other Tales* (1888) and *A House of Pomegranates* (1891).

In the final decade of his life, Wilde wrote and published nearly all of his major work (Beckson, 2021: 2). Wilde's first literary success came with his only novel, *The Picture of Dorian Gray*, published in *Lippincott's Magazine* in 1890, but it received quite a negative response from Victorian critics, mostly because of Basil Hallward's homoerotic feelings for Dorian Gray. The novel was released in book form in 1891, and this expanded and revised version toned down this aspect. However, Wilde's greatest talent was for writing plays, and his most acclaimed comedies were *Lady Windermere's Fan* (1892), *A Woman of No Importance* (1893), *An Ideal Husband* (1895), and the greatest of them all was *The Importance of Being Earnest* (1895). The last-mentioned work exposes Victorian hypocrisies, and since its publication it has achieved the status of an all-time theatre classic.

Wilde's social and literary success abruptly ended in 1895 following his arrest. He was sent to prison because of his homosexual relationship with Lord Alfred Douglas,

otherwise known as 'Bosie', even though Wilde was already married to Constance Lloyd. His experience in prison inspired two of his greatest works, *The Ballad of Reading Gaol* (1897), in which he revealed his concern for inhumane prison conditions, and *De Profundis* (partly published posthumously in 1905), a long letter to 'Bosie' about their relationship. The sentence caused the financial ruin of Wilde's family. Indeed, when he was released from prison in 1897, he went to Paris, and he spent his last three years in poverty, bad health, and exile from social and artistic life, supported by his friends' money. Wilde died of meningitis in 1900, alone, and in a cheap Parisian hotel.

As previously stated, *The Picture of Dorian Gray* is Wilde's only novel, as well as being the work that best sums up his own aesthetic theories about a life of sensation and pleasure as the supreme form of art. The novel tells the story of Dorian Gray, a charming young man who pursues his own aesthetic goals and is unaffected by age and excess, but his hidden portrait is disfigured by every detail of Dorian's immoral and hedonistic life.
The novel opens with the corrupt Lord Henry Wotton, a young aristocratic dandy, and Basil Hallward, a painter, contemplating the latter's unfinished portrait of a fascinating young man. Basil reveals that the sitter is Dorian Gray, reminiscing their first encounter and confessing that the painting will not be exhibited due to the fact that he fears that he has put too much of his own self into it. Then Dorian arrives and he is fascinated by Lord Henry's witticism about art and life. However, after hearing Wotton's warning that his youthful beauty will fade, Dorian has an extreme response to the finalized portrait, and being terrified by the idea of aging, he wishes he could maintain his youth and beauty while the painting alters with the passing of time. This wish is granted, and Dorian, influenced by Wotton, begins to dangerously seek every experience in life, leading an immoral and libertine life without altering his appearance, because the portrait is the one that records Dorian's sins and vices. For instance, as described by Parfitt (2013), Dorian "goes to parts of London that some people of his social stature never see, and finds a shabby theater", in which he meets and falls in love with actress Sibyl Vane, who then commits suicide following Dorian's decision to call of their engagement.
Despite being terrified by the changes on the portrait, Dorian still leads a decadent life. His reputation sours, but he is so charming and wealthy that he is still welcome in the highest circles (Parfitt, 2013). Dorian's immoral and criminal acts include murders,

blackmails, and corruption, too. Realising that confessing what he has done will absolve him of wrongdoing, Dorian decides to destroy the picture, but in doing so, he kills himself. After his death, the portrait resumes its original state, while the signs of age and physical corruption appear on Dorian's body.

2.2 Misogyny and representations of women in *The Picture of Dorian Gray*

This subchapter is dedicated to a specific aspect present in Wilde's novel, which is misogyny. Therefore, I give some examples of how women are insulted, degraded, and objectified in many different passages of the novel. In order to do so, in this part of the chapter I illustrate the social background of Wilde's own society, since his purpose is to reflect it in his literary works.

As presented by Buzwell (2014: 1), Wilde's *The Picture of Dorian Gray* is "a representation of how fin-de-siècle literature explored the darkest recesses of Victorian society and the often disturbing private desires that lurked behind acceptable public faces". As a matter of fact, the Victorian period was characterized by extreme gender segregation and sexism, which, according to Mohammed and Abdullah (2019: 533-534), "was reinforced through images and texts that implicitly and explicitly stated that work was manly, and therefore not appropriate for women". The dominant public sphere of work was reserved for men, and it was separated from the private, feminine sphere of the household. This shows that women were looked upon as 'the sex', which stressed that they were only identifiable by their natural biological functions and that their character was dominated by traditional gender roles (Parker, 1995: 6). While women's sexuality was denied, and their only passions linked to maternal instincts, men were full of desire, and not reprimanded when visiting prostitutes to fulfil their needs (Kersten, 2014: 11).

Thus, the highly criticised novel reflects the Victorian sexist nature inflected on women throughout that period. McKenna (2004) is in no doubt that not only was the much-loved Victorian raconteur a misogynist but that *The Picture of Dorian Gray* exemplifies how, in Oscar's world, art imitated life. In The Picture of Dorian Gray, women are portrayed as superficial, irrational and inferior in intellect compared to men (Wang, 2016: 2).

Wilde's work also presents several hate speeches against women, sexist comments, and misogynistic insults, mainly delivered by Lord Henry Wotton.

In the novel *The Picture of Dorian Gray*, one of the ways in which the author emphasizes misogynistic views is by choosing male main characters, such as Dorian Gray, Basil Hallward, and Lord Henry Wotton. Subsequently, the author outcasts women in his work in the same way that men treated women in real life during the Victorian Era. Indeed, the female characters in the novel are secondary, and, as Mohammed and Abdullah (2019) noticed, "they never play an effective part in the development of the actions of the novel". Sibyl Vane is the only female character who is relatively important to the story; however, she is still perceived as less important and inferior. Of course, there are other female characters, but they are mentioned merely because they are family members of the main male characters. For instance, Lady Agatha and Victoria Wotton, who respectively are Lord Henry Wotton's aunt and wife, do not have much to say, but the men do talk about them, and as may be expected, the content of their remarks is even more relevant, since they are downgrading comments against female characters or women in general.

The very first time a woman is mentioned in the novel is when Basil describes Lady Brandon and her "curiously shrill voice", and Lord Henry immediately replies that "she is a peacock in everything but beauty" (Wilde, 1891: 10). The two continue with further statements, such as "Lady Brandon treats her guests exactly as an auctioneer treats his goods" and "she tried to found a salon, and only succeeded in opening a restaurant" (Wilde, 1981: 11). Moreover, the next passage where a woman is mentioned is when Lord Henry states that "women have no appreciation of good looks", indirectly alluding to the fact that women's opinions are incorrect. However, one of the most remarkable quotes on women in the entire work is uttered by Lord Henry after Dorian has described Sibyl as a genius:

> My dear boy, no woman is a genius. Women are a decorative sex. They never have anything to say, but they say it charmingly. Women represent the triumph of matter over mind, just as men represent the triumph of mind over morals.
>
> (Wilde, 1891: 46)

The previous quotes about Lady Brandon might be personal, suggesting that Basil and Lord Henry may simply think of her as neither too nice nor pleasant, but this quotation – as well as other countless instances – undoubtedly exposes "how the men in the novel think about women in general", as Kersten (2014: 12) points out. This statement clearly reflects women's inferiority and discrimination in the Victorian age. Yet another instance of extreme sexism that emphasizes the social oppression that Victorian women were subjected to is once again delivered by Lord Henry:

> I am afraid that women appreciate cruelty, downright cruelty, more than anything else. They have wonderfully primitive instincts. We have emancipated, but they remain slaves looking for their masters all the same. They love being dominated.
>
> (Wilde, 1981: 95)

Lord Henry, who therefore imposes himself as a misogynistic man throughout the novel, describes women as primitive and weak, and, even worse, as slaves. Henry, who is an aristocratic upper class man, believes that women are naturally and biologically fit to be slaves and that it is men who can emancipate them from this slavery yet they still need male domination (Mohammed and Abdullah, 2019: 535). These words and insults are employed to convey the idea of male domination and women's inferior position. There are other countless passages in which Lord Henry discredits and disrespects women, one of them being:

> Never trust a woman who wears mauve, whatever her age may be, or a woman over thirty-five who is fond of pink ribbons. It always means that they have a history.
>
> (Wilde, 1981: 94)

This is an insult to Victorian women for being promiscuous and have illegitimate relationships out of marriage. However, there is not a similar standard for men, illustrating the inequality between the sexes once again (Kersten, 2014: 13). The female characters in the novel are always humiliated and shamed for their so-called flaws, but they are never praised for their personalities and intellects.

In addition, women are often judged for their appearances and looks, as well as for their age. As Mohammed and Abdullah (2019: 539) noticed, "the most prominent examples are pronounced by Lord Henry", and in one of his monologues he states that:

> I find that, ultimately, there are only two kinds of women, the plain and the coloured. The plain women are very useful. If you want to gain a reputation for respectability, you have merely to take them down to supper. The other women are very charming. They commit one mistake, however. They paint in order to try and look young. Our grandmothers painted in order to try and talk brilliantly. *Rouge* and *esprit* used to go together. That is all over now. As long as a woman can look ten years younger than her own daughter, she is perfectly satisfied. As for conversation, there are only five women in London worth talking to, and two of these can't be admitted into decent society.
>
> (Wilde, 1891: 46)

Lord Henry's statements and opinions are important because they demonstrate the fact that Victorian men never found any real interest in women, solely seeing them as a distraction or company. Plain, ordinary women are convenient in order to be respected in high society, whereas the other category of women identified by Lord Henry draws negative attention, and hence it is better to avoid being seen in public with them.

The other character that often shows a misogynistic point of view is Dorian Gray, even though he is not as explicit as Lord Henry. Nonetheless, Dorian's remarks and insults are still hurtful and degrading. The following lines exemplify Dorian's thoughts on women:

> Ordinary women never appeal to one's imagination. They are limited to their century. No glamour ever transfigures them. One knows their minds as easily as one knows their bonnets. One can always find them. There is no mystery in any of them. They ride in the park in the morning and chatter at tea parties in the afternoon. They have their stereotyped smile and their fashionable manner. They are quite obvious.
>
> (Wilde, 1891: 49)

As Mohammed and Abdullah (2019: 535) observed, Dorian seems to think that women are "shallow and trivial", "busy with their appearances", and "easily approachable and hence controllable". He blames women for being predictable and stereotypical ignoring that they acquired such traits as a result of the oppression and discrimination imposed on women. On the other hand, men are seen as superior since they made the rules, and they are the ones who decide what is decent for women and what is not.

Dorian is clearly influenced by Lord Henry, and this reflects the sexism present in Victorian society. As a matter of fact, Dorian was originally presented as an innocent young man who respects and love women, but his opinion changes to match the prevailing sexism and misogyny that characterize his society.

Mohammed and Abdullah (2019: 535) highlight that "the sexist nature of the novel is also traced in the domain of names, surnames and titles". A case in point is Sibyl Vane, whose surname is pronounced as "vain". Her name seems to be an indicator of the vain fate of women in Victorian society.

Furthermore, according to the traditions of that period, women would change their last name after marriage, adopting their husband's family name. Taking the husband's surname coincided with the appropriation by the husband of the wife's possessions and property (Mills, 1995: 61-62). This form of possession and objectification is also present in the novel. For instance, the first name of Dorian's housekeeper, Mrs Leaf, is not stated, since she is considered a possession of her husband and hence, she is not made known by her own name. Similarly, Lady Brandon and Lady Henry, who are mainly identified by their husbands' names. In *The picture of Dorian Gray* women, as stated by Mohammed and Abdullah (2019: 536), "are depicted as nameless characters whose lack of identity is only complete when it is bound with male identity".

In *The Picture of Dorian Gray*, women are portrayed as superficial, irrational and inferior in intellect compared to men (Wang, 2016: 2). Since Wilde's work reflects the misogynistic, sexist attitude typical of the Victorian society, men rule the novel. Subsequently, the female characters who are mentioned or are allowed to speak are objectified, insulted, and belittled.

2.3. Overview of corpus stylistic analyses of Wilde's works

The final section of the second chapter of my work briefly describes some examples of corpus stylistic analyses. Both of them have similarities with the research I plan to carry out, since the two of them observe a selection of Wilde's literary works. I also highlight the methods involved in the respective investigations, and, I report the most significant conclusions.

The computer-assisted study of literary texts can make an important contribution to the study of linguistic devices by increasing the scale at which they are analyzed (Mahlberg, 2012: 52). As it is a literary text that remains relevant in today's society, *The Picture of Dorian Gray* is still subjected to analyses and research, but not many of them are corpus stylistic studies. For example, Nordentoft's analysis (2020) investigates Wilde's novel by way of corpus methods.

Nordentoft (2020) used four corpora in her investigation of Wilde's only novel: *The Picture of Dorian Gray*, which is the target text; a corpus of Oscar Wilde's collected works; a Gothic corpus, which consists of similar texts in terms of genre, but not of period; and a corpus of texts that are contemporary to Wilde's one. In addition, in this analysis authors are not represented more than once in the same corpus, with the aim to provide greater balance.

In the corpus-driven analysis, Nordentoft (2020: 54) conducted "a comprehensive counting of the tags in all corpora, in order to be able to locate elements that are significantly more frequent (or infrequent)" in *The Picture of Dorian Gray* than in the three reference corpora. Referring to their work, this counting disclosed a relative underuse of nouns (the NN-tag) in the target text of this specific work. Furthermore, probably due to the countless dialogues present in the novel, personal pronouns (PP) are highly overused, since Biber and al. (1999) assess that pronouns are used much more frequently in conversations and common-speech than in writing. There also seems to be a pattern regarding coordinating conjunctions (CC) and prepositions/subordinate conjunctions (IN) and general joiners (:) (Nordentoft, 2020: 57). Such word types are remarkably underused in Wilde's novel, and Nordentoft (2020: 57) concludes that this "may indicate that the text has a low degree of explicit cohesion". This characteristic

could be linked to the significant overuse of sentences (SENT) in the target text, meaning that its sentences tend to be shorter than the ones in the three reference corpora.
As regards word classes, Nordentoft (2020) noticed that both adjectives and adverbs, which indicate a degree of descriptive language, are rather revealing in terms of style. Indeed, adverbs and adjectives seem to be more frequent in *The Picture of Dorian Gray* than in the Gothic corpus, but not compared to the remaining reference corpora. This is the reason why it is not possible to make generalizations regarding the degree of descriptive language in the target text.

Another way to approach corpus-driven analysis is looking for keywords, as used by both Mahlberg (2012) and Fischer-Starcke (2010). This measure determines how key a word is compared to the reference corpus. Main characters' names and words that are specific to the theme of a novel are usually keywords, since they characterize a text. For instance, Nordentoft (2020: 59) highlighted that 'portrait', 'picture', and 'canvas' are "highly key compared to all reference corpora".
Moreover, verbs like 'seemed', 'murmured', and 'cried' are significantly more frequent in the target text than in the reference corpora. Nordentoft (2020: 95) argued that "the verb 'seem' points to an *impressionist* mode of narrative, in which the characters' subjective understanding of the world was important". The fact that inner experiences are found to be central is matched with what the researcher found in the corpus-assisted analysis, that is the presence of a "*poetic* mode in the novel, since impressionism and poetry share similarities and can be linked together", according to Nordentoft (2020: 95). Therefore, both analyses displayed the aesthetic-poetic quality of Wilde's *The Picture of Dorian Gray*. Also, the co-occurrence of 'murmur' and 'cry' adds "a dramatic effect and an effect of volume", as defined by Nordentoft (2020: 96).

Another example of a corpus-aided study is Junmei's one (2017), which observed the language features of Wilde's masterpiece *The Importance of Being Earnest* (1895). The corpora used for the research were the text of the play and the British National Corpus.
One of the major findings of this study is that, in Junmei's words (2017: 37), "Wilde holds a disapproving attitude against the society by frequently using the words of negativity or in the negative context". For instance, adjectives like 'heartless', 'horrid', 'unfortunate', 'wicked', 'vulgar', and 'absurd' are significantly recurring throughout the

considered text. In addition, as concerns adverbs, Junmei (2017) noted that their main function is describing negative qualities, and many of them are used as positive modifiers. For example, the adverb 'perfectly' is frequently employed in negative contexts. This finding seems to suggest that Wilde is implicitly criticizing social conventions. All the lexical features are showing Wilde's ironic attitude to the world he was living (Junmei, 2017: 37). Indeed, the artificiality of Wilde's style used in this work reflects the sophisticated society typical of the Victorian period.

Furthermore, Junmei (2017) examined the syntactical features of the comedy, and he noticed that the number and position of negatives is significantly relevant. Not only is the word 'not' repeated several times, but there are also other words and expressions that appear in the text of the play that point towards the same directions, such as 'don't', 'isn't', 'won't', 'can't', 'never', and 'hardly'.

On the other hand, Binongo and Smith (1999) applied statistical techniques to literature and linguistic studies, focusing on Oscar Wilde's works. They used the Cartesian plane to investigate "whether or not differences between his plays and essays can be expressed graphically" (1999: 781).

As regards the analysis of Wilde's plays, the researchers downloaded five of them that were available at Project Gutenberg: *The Duchess of Padua* (1883), *A Woman of No Importance* (1892), *An Ideal Husband* (1893), *The Importance of Being Earnest* (1894), and *Lady Windermere's Fan* (during or before 1892). After editing so that only the actual texts remained, the sentences written in foreign language were deleted. Subsequently, the texts were divided into blocks of 3000 words, and then a computer program counted the number of occurrences of 'of' and 'to', the two most common prepositions in the English language. Due to the fact that prepositions are not inflected and are least dependent on context, Binongo and Smith (1999: 782) affirmed that "their numbers potentially provide a reliable indicator of stylistic idiosyncrasies of authors". The analysts noticed that "although there is no essential difference in the usage of the preposition 'of' in the plays, the usage of 'to' in *The Duchess of Padua* is strikingly less than the mean" (1999: 783). Among the analysed plays, *The Duchess of Padua* is the earliest and the only one set in the second half of the 16th century. These characteristics might be the reasons behind these results.

On the other hand, Wilde's essays involved in the analysis are compiled in the following publications:

- *Intentions* (1891): *The truth of masks* (originally *Shakespeare and stage costume*, 1885), *The critic as artist* (originally *The true function and value of criticism*, 1890), *The decay of lying* (1889), and *Pen, pencil and poison* (1889).
- *Collected Works* (1908): *The rise of historical criticism* (1879), and the four lectures given in 1882, entitled *The English renaissance of art*, *House decoration*, *Art and the handicraftman*, and *Lecture to art students*.
- *The Soul of Man under Socialism* (1891) and *De Profundis* (written in 1897 and published complete in 1962).

However, *House decoration*, *London models*, and the six *Poems in prose* compiled in *Collected Works* were excluded, since they had less than 3000 words. Also, Project Gutenberg only provided an excerpt of the original *De Profundis*. It is important to notice that works such as *The critic as artist* and *The decay of lying* are dialogues, *De Profundis* is a letter, and *The Soul of Man under Socialism* and *The rise of historical criticism* are formal essays, thus, as highlighted by Binongo and Smith (1999: 783), "there are differences of genre among these works". A decreasing use of 'of' as Wilde aged is what emerged from this research.

Bringing together Wilde's plays and essays, the difference in genre is more predominant than other factors. The prepositions 'of' and 'to' certainly reflect the difference in genre, but they need assistance (Binongo and Smith, 1999: 785). At the same time, including a larger number of prepositions would preclude a two-dimensional scattergram, and the graphs would be more difficult to interpret with a dimensionality beyond three. Therefore, a technique to reduce the dimensionality, even a large one, to two was needed, and when it was subsequently expanded, it was shown that the 25 most frequent prepositions in the corpus of Wilde's plays and essays were: 'of', 'to', 'in', 'for', 'with', 'but', 'on', 'at', 'by', 'from', 'out', 'into', 'about', 'before', 'through', 'up', 'upon', 'over', 'between', 'after', 'without', 'against', 'down', 'under', and 'except'.

The analyses mentioned above are only some of the countless instances of corpus stylistic research. However, I find them to be of interest for my own study, since I plan to conduct

a corpus stylistic analysis of Wilde's *The Picture of Dorian Gray*, which I will do in the following and final chapter of my dissertation.

Chapter 3

My corpus stylistic analysis of Wilde's *The Picture of Dorian Gray*

The third and final chapter of the present thesis focuses on the corpus stylistic analysis that I conducted. In the first section of this chapter, I introduce both the corpora and the tools that I chose to use in my analysis, as well as my expectations in terms of findings. Then, in the following part, I proceed by presenting the results that emerged from the research. I conclude the third chapter by comparing the results with my initial expectations, reflecting on which of them were confirmed and which of them led to surprising, unanticipated results.

3.1. Introduction of my analysis: corpora, tools, and expectations

This fragment focuses on the corpora and tools that I decided to use to carry out my analysis. I present the texts included in the corpora, I explain how the concordancer *AntConc* works, and I also describe its main functions, which I used to analyse *The Picture of Dorian Gray*. At the end of this subchapter, I expose what I plan to search, and I also share my expectations as regards the findings.

3.1.1. Corpora

Before tackling the process of corpus design, corpus designers have to pose a number of questions related to the size, range of language varieties, the period, whether the corpora will be written or spoken and the number of the textual materials to be included in the corpora (Pearson, 1998: 50). The process of corpus design is not simply a choice of the collection of texts to be considered in the analysis, but, as Hussein (2017: 45) highlights, "it is intended to show a language or some part of a language", and this is also linked to the purpose of the corpora.

In my work, two corpora are used: the target corpus is made up by *The Picture of Dorian Gray*, and the reference corpus comprehends other works written by Oscar Wilde (henceforth *Wilde*). The reference corpus is required to test the validity of my results, since in order to state that a certain pattern is significant in the target corpus, I must check

that it is not common in the reference corpus. They are all specialized corpora, as all of them are designed to represent special types of language varieties (McEnery, Xiao & Tono, 2006: 15).

My analysis includes only texts retrieved from *Project Gutenberg*. This is a weakness in terms of methodology, but at the same time previous methodologically similar works pose the same issue. The target corpus, *The Picture of Dorian Gray*, consists of 78,866 word tokens, whereas *Wilde* consists of 337,628 word tokens, and it is made up by 15 texts (see Table 1).

Title	**Text type**	**Tokens**
A House of Pomegranates	Collection of fairy tales	33,238
A Woman of No Importance	Play	22,117
An Ideal Husband	Play	30,514
De Profundis	Letter	17,683
Impressions of America	Travel description	4,090
Intentions	Collection of essays	58,204
Lady Windermere's Fan	Play	19,391
Poems with the Ballad of Reading Gaol	Poetry	37,033
Salomé: A Tragedy in One Act	Play	11,541
The Canterville Ghost	Short story	11,304
The Duchess of Padua	Play	22,262
The Happy Prince and Other Stories	Collection of fairy tales	16,195
The Importance of Being Earnest	Play	20,710
The Soul of Man Under Socialism	Essay	14,352
Vera; Or the Nihilists	Play	18,994

	Total:	**337,628**

Table 1. *Wilde* corpus contents.

3.1.2. Methodology

In modern corpus linguistics and stylistics, the use of computer software is both pivotal and instrumental (Nordentoft, 2020: 43). The corpus stylistic analysis is conducted using the concordance software, *AntConc*, provided by Laurence Anthony, professor of Educational Technology and Applied Linguistics at Waseda University (Anthony, 2015). *AntConc* is a concordancer, a programme that can examine a corpus and its elements, such as single words, frequent clusters of words (also called n-grams), lemmas, which words they appear with in their co-text, collocates, and more. In other words, as Stubbs (2014: 53) remarks, the software can find "the frequency of linguistic units (words, phrases, and lexicogrammatical structures) and their distribution (within an individual text or across a corpus)".

With the help of the corpora and the tools that I have previously described, I intend to conduct a corpus stylistic analysis of *The Picture of Dorian Gray*. More specifically, the goal of my investigation is to identify the frequent terms and sequences of words that convey misogynistic messages in Wilde's novel. However, since my intention is to conduct a corpus-assisted analysis, which tests *a priori* assumptions before starting the proper analysis, I would like to illustrate my own personal expectations on the future findings, which I plan to subsequently compare and contrast with the actual results.
Firstly, I assume that the words 'woman' and 'women' are mentioned several times in the target corpus, because when I read the novel for the first time, I had the impression that women are a frequent topic of conversation, especially in Dorian Gray and Lord Henry's dialogues. In addition, since many of these conversations have to do with physical appearance, beauty, and charm, and many of them are referred specifically to women, I believe that adjectives that describe these qualities tend to be recurrent throughout the novel. The terms 'charm' or 'charming' themselves, maybe even employed with negatives such as 'isn't' or 'not', might have been used consistently in Wilde's novel, probably even in passages in which women are criticized, mocked, or insulted. Indeed,

more generally, I presume that negative words and negative contexts are often present in hate speech towards women. In addition, I expect to find several words and expressions that point to the objectification of women, who are merely described as objects, accessories, or just as individuals that must entertain and keep company to men. I expect to find this negativity and attitude towards the feminine side due to the fact that in Victorian society women were seen and portrayed as powerless and inferior, and they often were the victims of careless treatment from men, which is yet another aspect that emerges in the novel.

3.2. Results

In the second section of the third chapter of my thesis, I explain how I conducted my analysis thanks to the tools and corpora that I mentioned in the previous subchapter. Therefore, the following pages show and discuss the results that emerged from the research. In addition, since the corpora used in my investigation are of different sizes, I normalize the results of my analysis, which means converting each frequency into a value per thousand words. In this way, the results can be compared more accurately and realistically.

I started my analysis by using the concordance tool on *AntConc*, searching for the words 'woman' and 'women' in the target corpus in order to find out how many times these words appear in the novel and in what kind of patterns. As expected, a number of results were found (respectively 29 and 49 hits), as Figure 1 and Figure 2 show.

.

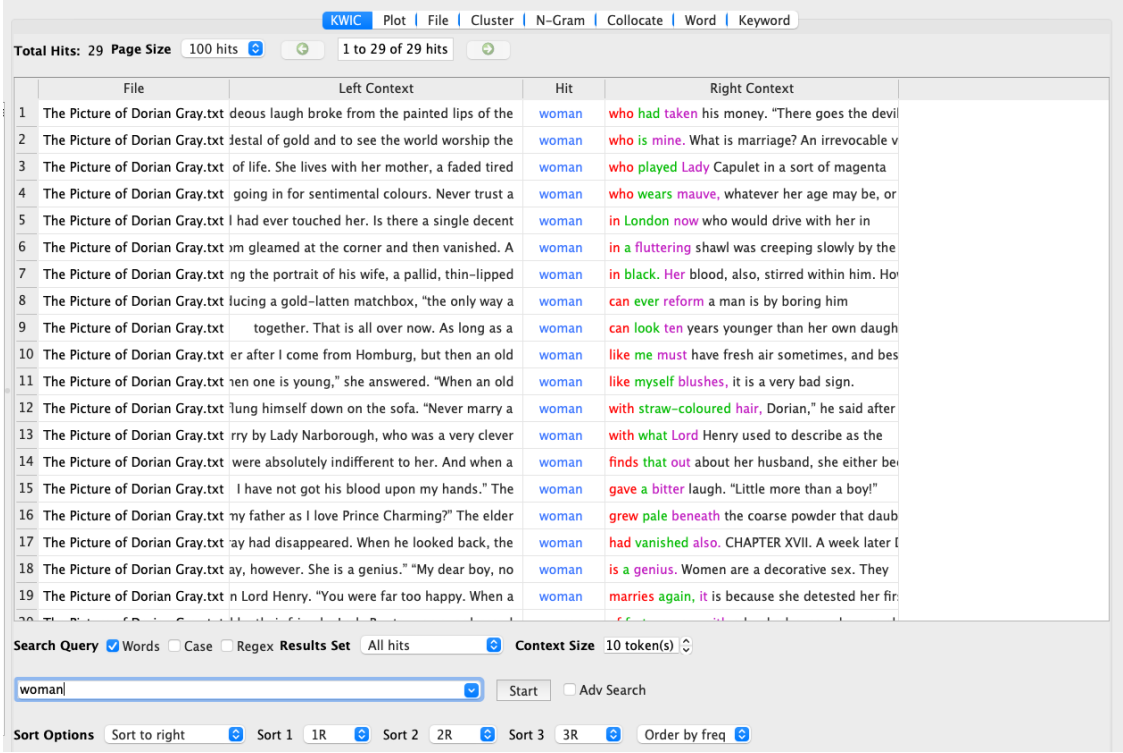


Figure 1. Hits for the word 'woman' in *The Picture of Dorian Gray*.

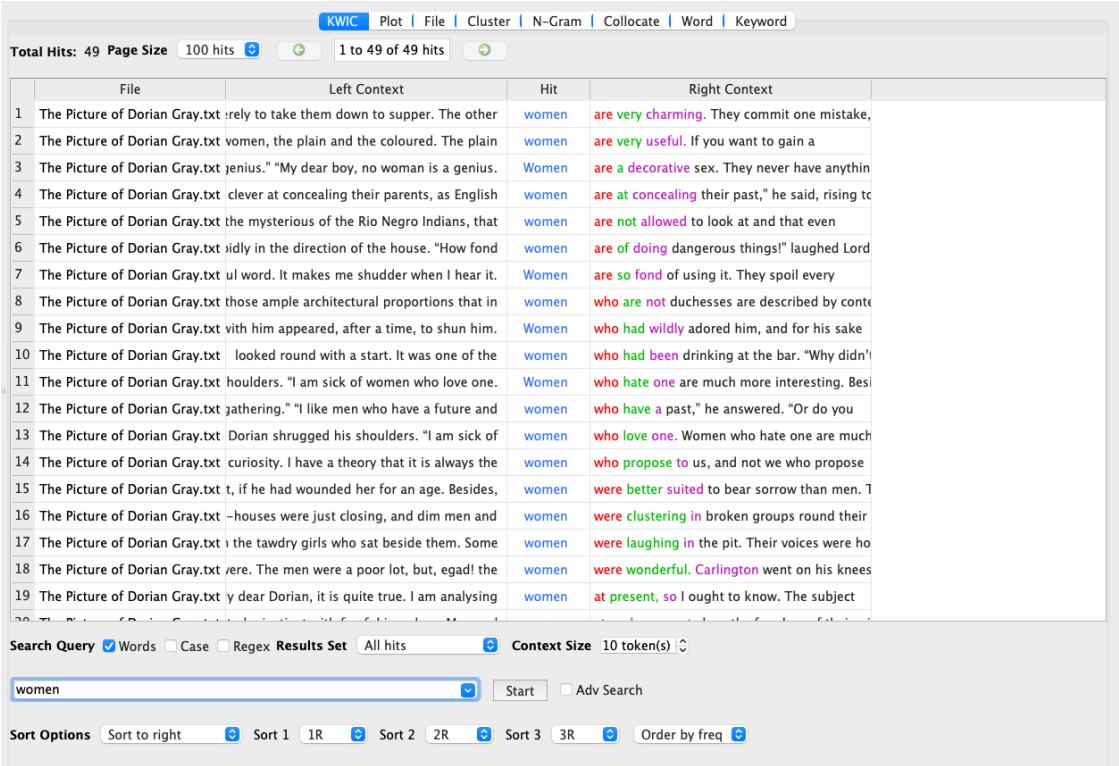


Figure 2. Hits for the word 'women' in *The Picture of Dorian Gray*.

What already emerges from Figure 1 and Figure 2 is that both of the terms ‘woman’ and ‘women’ are often found with negative co-text. In most of these lines, women are judged for their appearance and age, described in a misogynistic way, judged for their past, and their mind and intelligence are belittled. For instance, in Figure 1, line 2, “the woman who is mine”, denotes a sense of possession. Then, in line 4 it is also suggested to “never trust a woman who wears mauve”, and another example suggests that a woman should “look ten years younger than her own daughter” (see line 9 in Figure 1). However, line 3 is the strongest one among the results shown in Figure 2, and it states that “no woman is a genius”, and women are described as “a decorative sex. They never have anything to say, but they say it charmingly”. These examples are interesting and relevant, because they show women inferiority and their role and perception in Victorian society, which is reflected in Wilde’s novel.

Another similar aspect that becomes apparent by observing the noun ‘woman’ has to do with collocates, which are words that typically co-occur with other words. Indeed, the search, which was sorted by frequency, showed that, besides the article ‘a’, and the pronoun ‘who’, the word ‘woman’ is preceded by the adjective ‘elder’ twice in Wilde’s novel (see Figure 3). This specific adjective, even if it only occurs a very limited amount of times, still seems to point towards the same direction, meaning that a woman’s age is not an overlooked matter, and, as discussed above, old age does not usually have a positive connotation in *The Picture of Dorian Gray*.

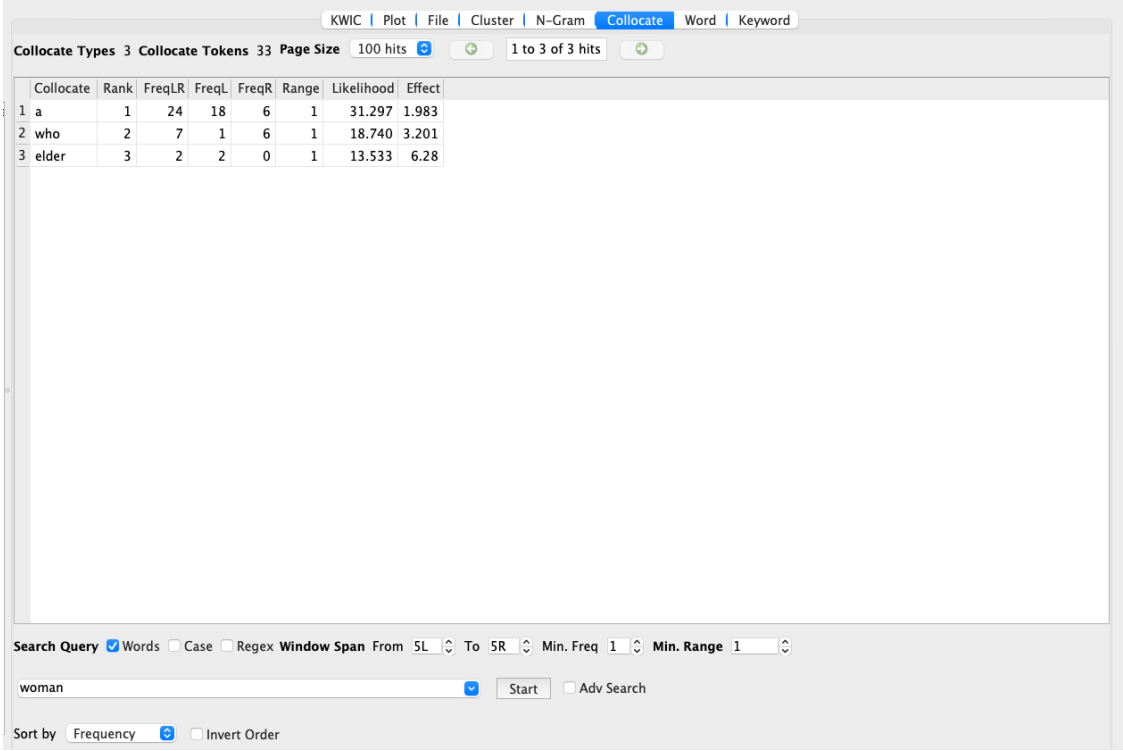


Figure 3. Collocates of the word ‘woman’ in *The Picture of Dorian Gray*.

In the next step of my analysis, I conducted the same search in the reference corpus, *Wilde*. Figure 4 and Figure 5 respectively show that the words ‘woman’ and ‘women’ are frequent in this corpus, too. What can be observed from the results is that some of them are also misogynistic or unflattering remarks towards women, such as “a mad woman who is seeking everywhere for lovers” (see line 2 in Figure 4), “a woman who has sinned should be punished” (see line 10 in Figure 4), and in Figure 5, line 10, “women are not meant to judge us”, which is uttered by a male character.

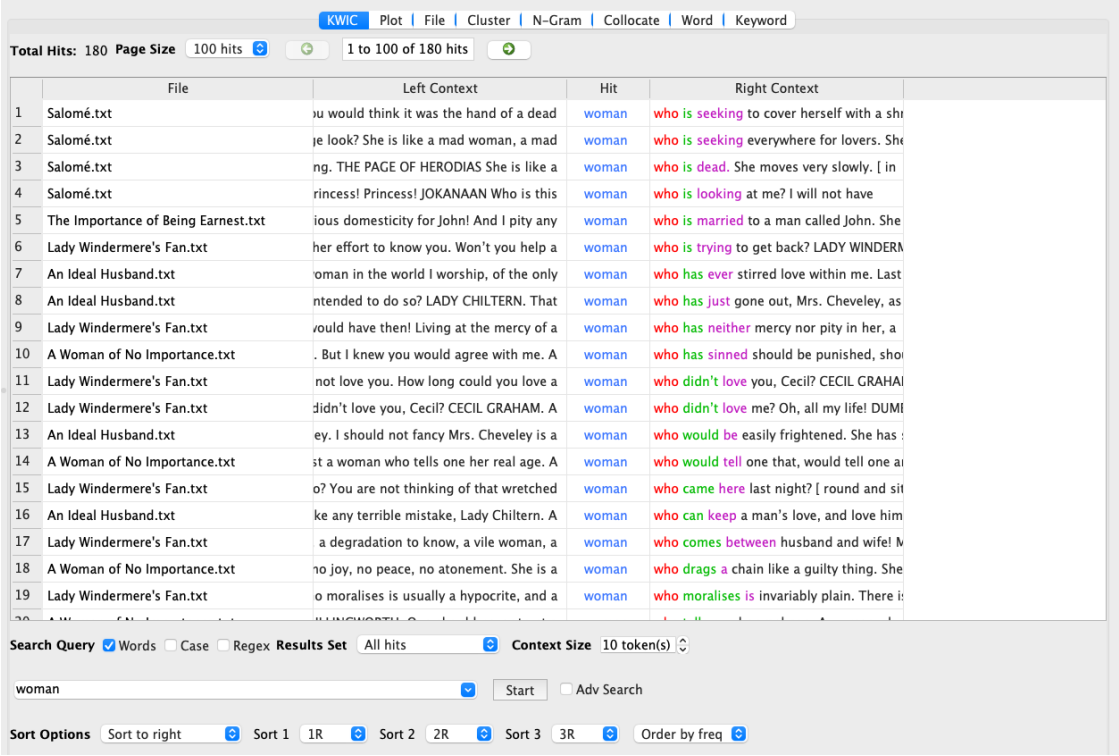


Figure 4. Hits for the word 'woman' in *Wilde*.

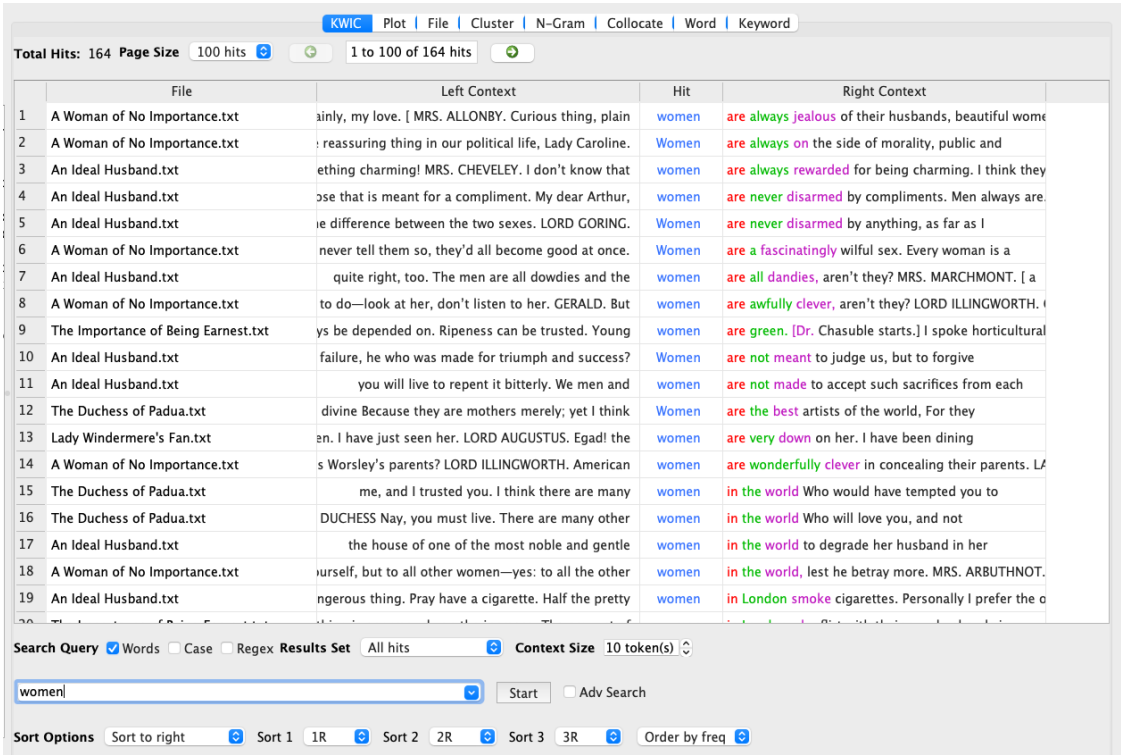


Figure 5. Hits for the word 'women' in *Wilde*.

At this point, it might be tempting to simply conclude that the number of hits for the words 'woman' and 'women' in *Wilde* is higher. However, the reference corpus is bigger than the target corpus, and therefore there are more opportunities for these terms to occur. This is the reason why I normalised the results obtained in both *The Picture of Dorian Gray* and *Wilde*, so that it is possible to compare them in a more realistic way. In order to do this, I used the following formula:

$$\text{Frequency per thousand words} = (\text{frequency} \div \text{text no. words}) \times 1{,}000$$

In the target corpus, the normalised results were 0.36 for 'woman' and 0.62 for 'women', whereas in *Wilde*, the normalised results were respectively 0.53 and 0.48. Comparing these results, it is possible to notice that there are not huge differences. Hence, the terms 'woman' and 'women' are used more or less with a similar frequency in the target corpus and in the reference corpus.

Subsequently, I repeated the collocate analysis in the reference corpus to investigate whether 'woman' produces similar results even in *Wilde* (see Figure 6).

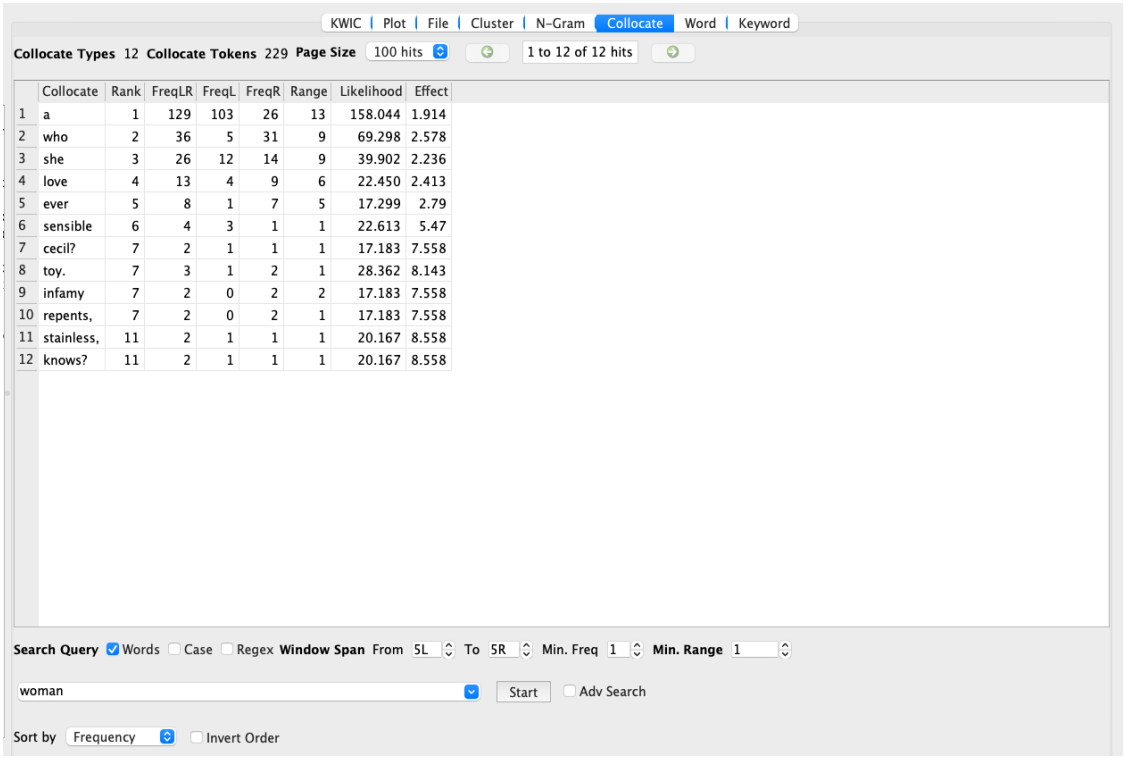


Figure 6. Collocates of the word 'woman' in *Wilde*.

As illustrated in Figure 6, the article 'a' (line 1) and the pronoun 'who' (line 2) are the most recurrent collocates in the reference corpus as well, but this is not relevant to the main goal of my search. On the other hand, it is curious to see that both of the terms 'love' and 'toy' appear in the same list (see respectively line 4 and line 8), because the former insinuates a more positive and romantic idea of women, whereas the latter suggests the opposite, that is objectification, inferiority, or even the idea of something that you can play with until you are done with it. Indeed, in this line, one of the characters of *A Woman of No Importance* (1893) expresses his fears that another male character might see women as a toy:

> Kelvil: I am afraid, too, that Lord Illingworth regards woman simply as a toy.
>
> (Wilde, 1893: 14)

I find this result to be quite interesting, because, since in *Wilde* there seem to be relatively more appreciative comments towards women than in the target corpus, this might mean that *The Picture of Dorian Gray* could be one of Wilde's most sexist works. Nonetheless, this could also be due to the fact that the reference corpus includes literary works that deal with various, dissimilar themes, and some of them might express different points of view.

For this collocate analysis, I reported only the findings of the singular noun 'woman' in the target corpus as well as in the reference corpus, but I actually repeated the same steps for its plural, 'women'. However, in both cases the search did not reveal any surprising results compared to the findings of 'woman', considering that they were similar to the ones previously discussed. For this reason, I decided not to report those results in this thesis, and I will instead move forward with the analysis of other terms.

Then, I carried out a search for the name 'Sibyl' in the target corpus because she is the only semi-prominent female figure in *The Picture of Dorian Gray*, and I wanted to observe which words collocate with her name. I do not expect to find the name 'Sibyl' in *Wilde*, since this is the only character in Wilde's works who bears the name Sibyl.

As predicted, ‘Sibyl’ is repeatedly mentioned throughout the novel, and Figure 7 shows that her name is associated with death and suicide (see lines 1, 2, and 12), which she commits at a certain point in the novel, but there are also allusions to love and marriage (for instance, see lines 3, 9, 15, and 19), and some praises and compliments as regards her acting skills are also present (see lines 5 and 6).

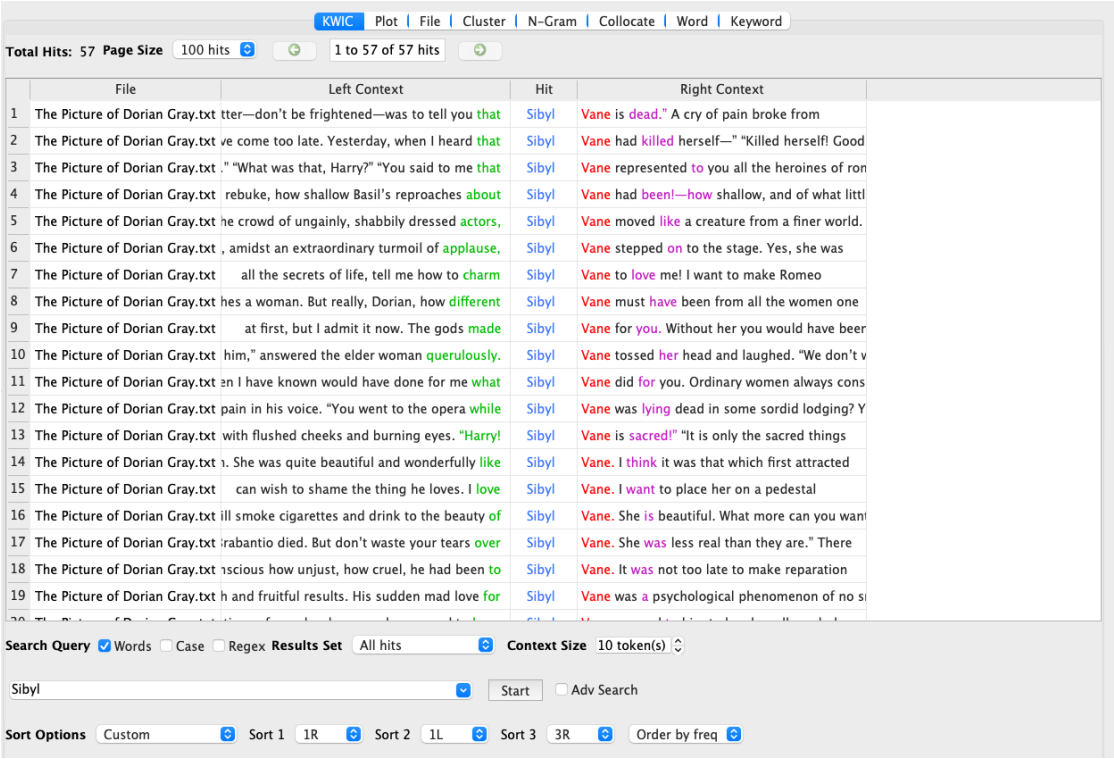
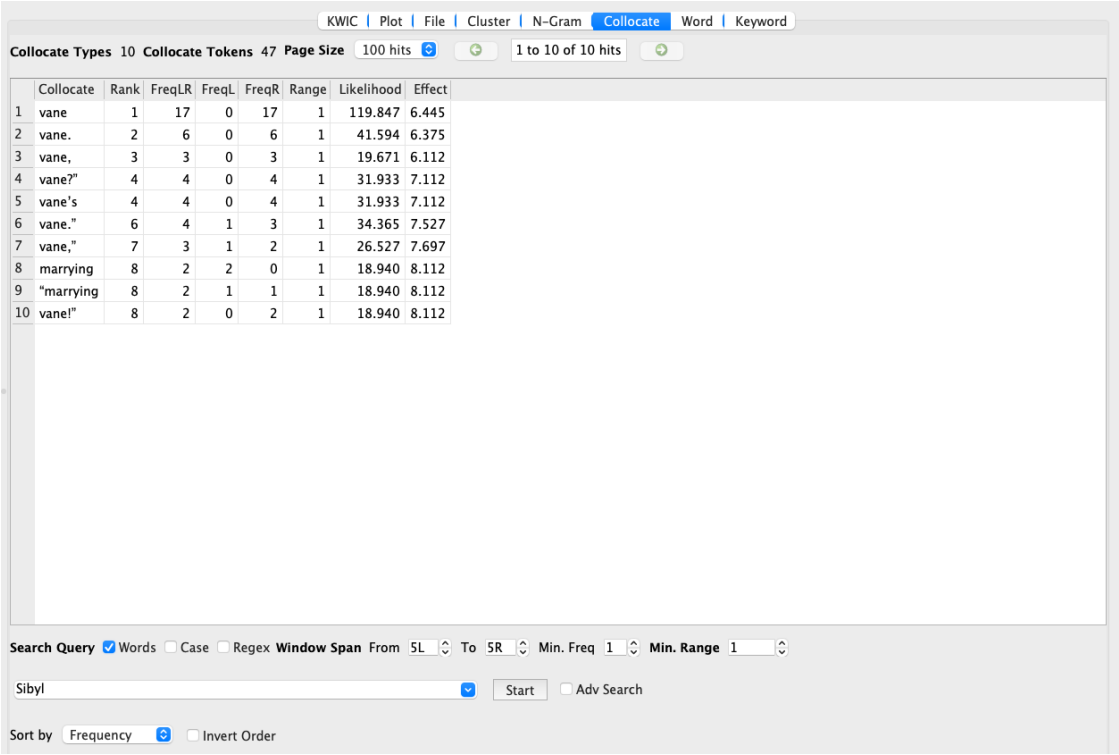


Figure 7. Hits for the name ‘Sibyl’ in *The Picture of Dorian Gray*.

Although the results in Figure 7 do not appear to be too misogynistic, it is important to remember that, as discussed in the previous chapter of the present thesis, at first Dorian Gray appreciates Sibyl and her talent, he even describes her as a genius. But later in the novel, also influenced by Lord Henry, who indirectly insults her and women more at large, Dorian drastically changes his mind, and thus he stops praising her, which indirectly led her to commit suicide.

In addition, due to the fact that Sibyl is often called and referred to with her complete name, the collocate analysis shows that her first name is very frequently followed by her surname, which is Vane (see lines 1-7 and line 10 in Figure 8). Another aspect that emerges from this search is that, in the target corpus, her name is accompanied twice by

the verb “marrying” (see lines 8 and 9 in Figure 8). This is curious, because Sibyl is mentioned in an appreciative speech in which she is praised for her wits and talents only at the beginning of the novel, that is only briefly and partially, but her name co-occurs more than once in lines that are related to the idea of marriage. However, marriage has a questionable patriarchal past, and for this reason the fact that we see Sibyl’s character mainly in relation to a much more important male character, Dorian Gray, is interesting and revealing. These findings are displayed in Figure 8.



	Collocate	Rank	FreqLR	FreqL	FreqR	Range	Likelihood	Effect
1	vane	1	17	0	17	1	119.847	6.445
2	vane.	2	6	0	6	1	41.594	6.375
3	vane,	3	3	0	3	1	19.671	6.112
4	vane?"	4	4	0	4	1	31.933	7.112
5	vane's	4	4	0	4	1	31.933	7.112
6	vane."	6	4	1	3	1	34.365	7.527
7	vane,"	7	3	1	2	1	26.527	7.697
8	marrying	8	2	2	0	1	18.940	8.112
9	"marrying	8	2	1	1	1	18.940	8.112
10	vane!"	8	2	0	2	1	18.940	8.112

Figure 8. Collocates of the name ‘Sibyl’ in *The Picture of Dorian Gray*.

The following stage of my analysis has to do with the term ‘charm*’ and its respective nominal, adjectival, adverbial, and verbal forms. In order to do so, I used a wildcard character, which Merriam-Webster’s Learner’s Dictionary defines as “a symbol (such as ‘?’ or ‘*’) that is used in a computer search to represent any letter or number”. Therefore, a wildcard stands for an unknown or unpredictable factor, it means that zero or more characters are following the previous letters. This feature is extremely useful because it leads to more results, as demonstrated in Figures 9 and 10: the search for ‘charm’ only has 20 hits, whereas the search for ‘charm*’ has 83, because the latter search includes

more forms in the results, such as the adjective ‘charming’, the term ‘charm’ used both as a noun and as a verb, and, of course, as a verb it can be present in all of its conjugations, for instance ‘charms’, ‘charmed’, ‘charming’, and so on.

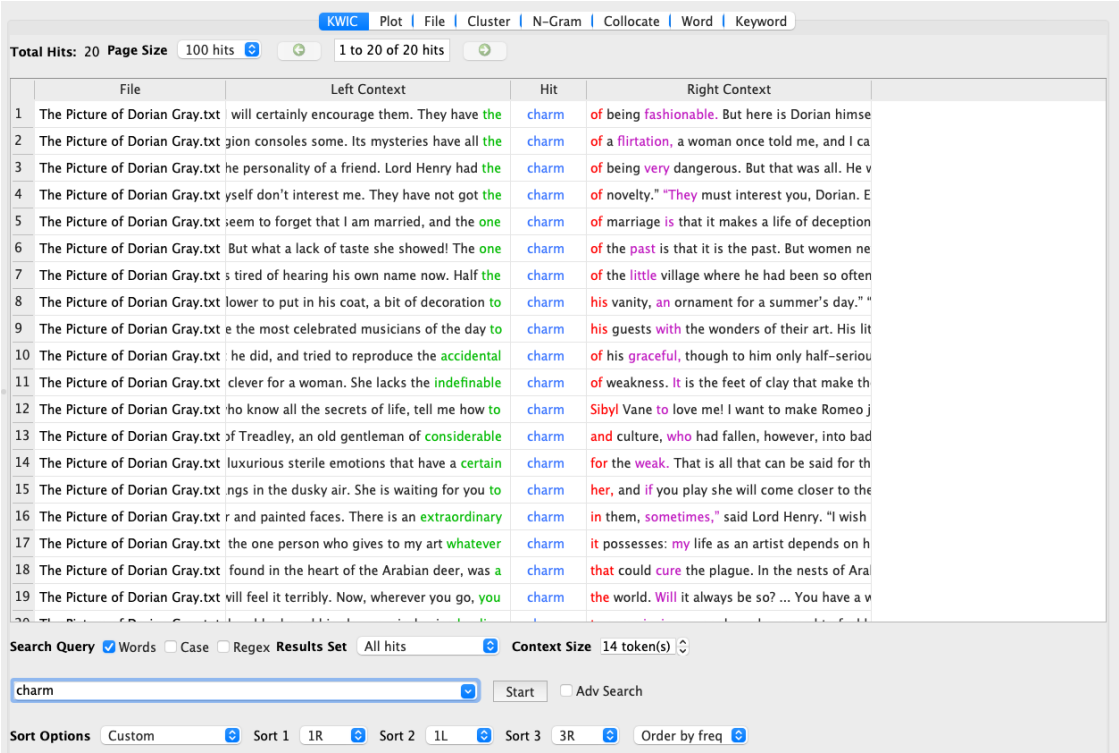


Figure 9. Hits for the word ‘charm’ in *The Picture of Dorian Gray*.

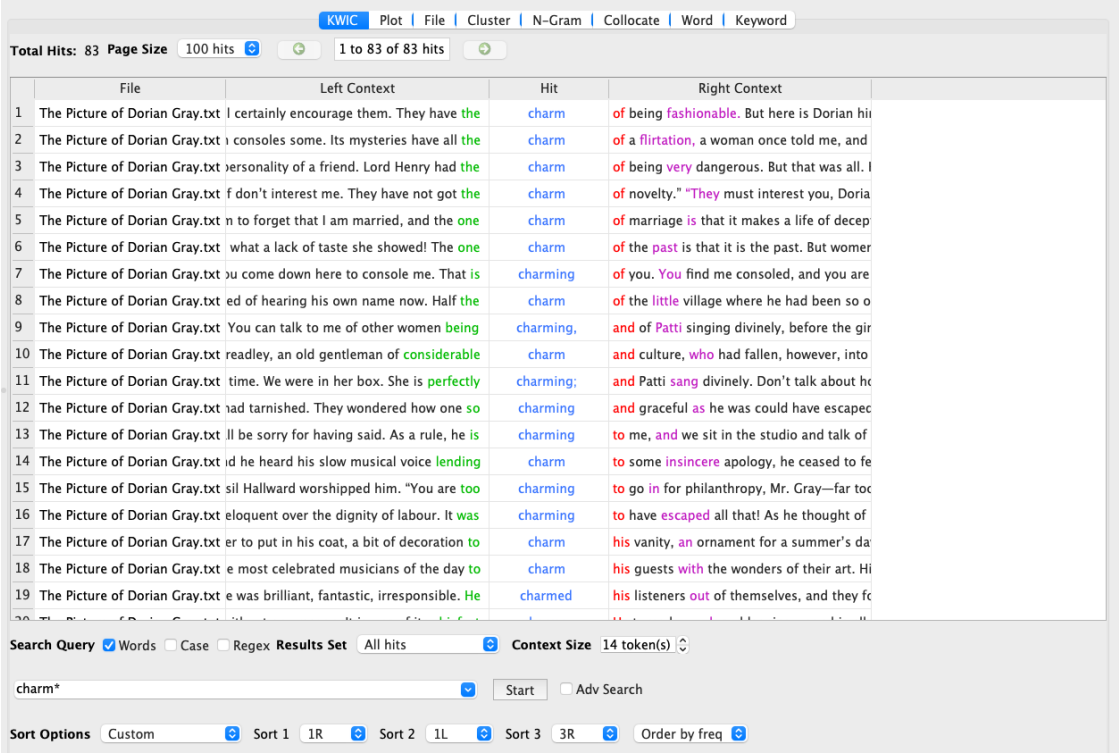


Figure 10. Hits for the word 'charm*' in *The Picture of Dorian Gray*.

Firstly, I typed 'charm*' in the search box and conducted this search in *The Picture of Dorian Gray*, and I chose this word in its many forms and conjugations because when I read the novel for the first time, I noticed that they are quite recurring, and I had the impression that they were mostly referred to the male characters, so I wanted to check if my assumptions were true. As demonstrated by Figure 10, my initial expectations were confirmed, since the number of hits is quite big (83 hits). The terms 'charm' or 'charming' are used a number of times, mainly to compliment or praise male characters (see lines 10, 12, and 15 in Figure 10), or to somehow allude to something that the male character who uttered these words finds interesting, pleasant, or enjoyable (for example, see lines 6 and 8 in Figure 10). Moreover, what can be observed by looking at the findings is that some of these results also have to do with women, but, once again, they are direct or indirect insults. Indeed, the analysed words, if referred to female characters, have a drastically different connotation. For instance, when Lord Henry complains about marriage and speaks ill of married life, the word 'charm' denotes a negative idea (see, for example, lines 1 and 5 in Figure 10).

In the target corpus, I also conducted a collocate analysis using a wildcard character. I searched for 'charm*', but there were no noteworthy findings, with the collocates of the words 'being', 'prince', and 'philanthropy'. These results do not lead to new, additional information, and therefore are not relevant to the purpose of the present analysis, and for this reason I decided to not report and discuss them in my dissertation.

The same search was carried out in the reference corpus. As illustrated in Figure 11, the search for 'charm*' revealed that in *Wilde* there are 116 hits. In order to make a realistic comparison between the corpora, I used the same formula as before to calculate the normalised results, which were 1.05 for the search for 'charm*' in the target corpus, and 0.34 in the reference corpus. In this case, the difference between the two corpora is outstanding, because it demonstrates that words like 'charm', 'charmed', 'charming', and so on, are indeed much more present in *The Picture of Dorian Gray* than in the other works of Wilde's.

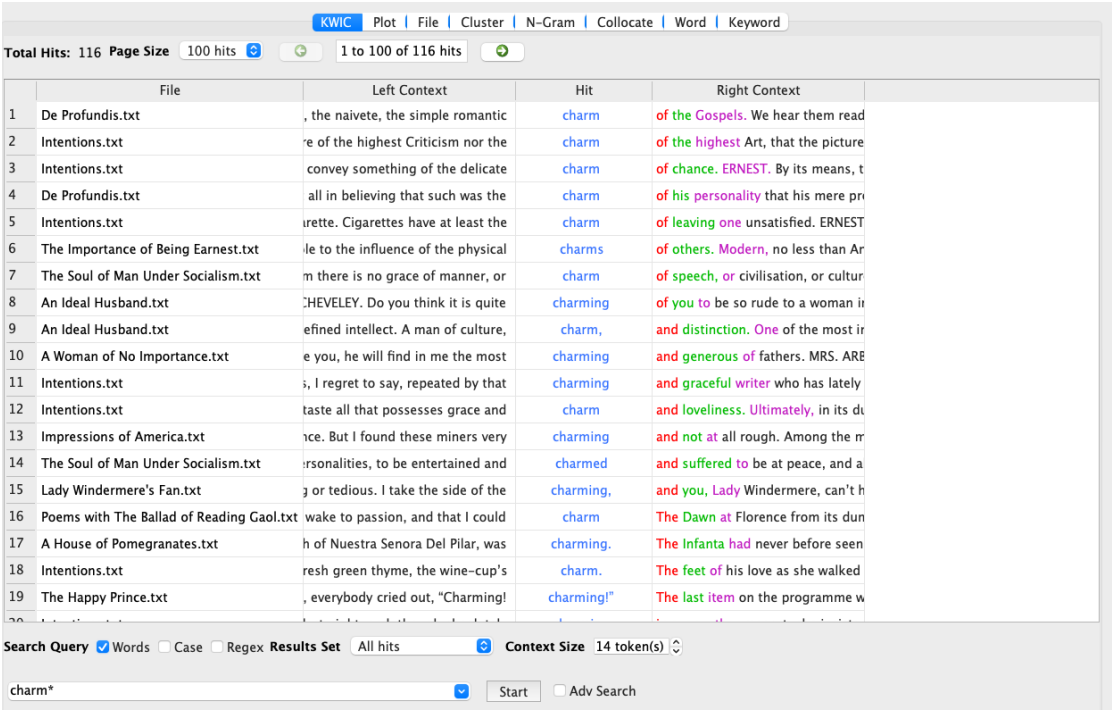


Figure 11. Hits for the word 'charm*' in *Wilde*.

In this corpus, the findings do not exactly point in the same direction, meaning that most of them do not refer to women. However, I noticed that there is a surprising result (see

line 8 in Figure 11) that highlighted a line of one of Wilde's plays, *An Ideal Husband* (1893):

> Mrs. Cheveley: Do you think it is quite charming of you to be so rude to a woman in your own house?
>
> (Wilde, 1893: 138)

What makes this line extremely interesting is the fact that it is uttered by a female character, Mrs Cheveley, who uses the term 'charming' not to compliment someone else in the play, or to say something pleasant, but as an attempt to stand up for herself against a male character, Lord Goring. Therefore, this finding has to do with women's empowerment, unlike most of the previous ones, because it is an example of how a woman can respond and react to disrespectful comments or behaviours. It is undeniable that Mrs Cheveley's character is the *femme fatale* of the play, but it is still interesting to analyse another trait of her personality, that is her refusal to conform to the role of a traditional wife.

The collocate earch for 'charm*' conducted in the reference corpus, similarly to the one carried out in the target corpus, did not disclose any particularly relevant outcome, considering that the only collocate was 'cigarettes', with a total of merely 2 hits. Since this finding is not pertinent to my work, I will move on to the following part of my analysis.

Other terms that I wanted to investigate in the target corpus and in the reference corpus are the ones related to external, physical beauty. In order to do so, I decided to search for 'beaut*', thus using once again a wildcard character. Since Dorian Gray's physical appearance is extremely important in the novel, I assumed that several results might refer to him, but not every single one of them.

Figure 12 and Figure 13 display respectively the concordances and the collocates disclosed by searching for 'beaut*' in the target corpus. This term has 101 hits (see Figure 12), and, as I expected, some of them are linked to the main character of the novel and to the concept of beauty in itself, while others refer to women. More specifically, Sibyl Vane is mentioned more than once, and at a certain point in the novel, the adjective 'beautiful' is employed to describe how Dorian's life would have been with her (see line 6 in Figure

12). Although these findings are not entirely related to misogyny, what can still be noticed is the fact that women do not play a big role in *The Picture of Dorian Gray*, and that, the few times that they are mentioned, they are either objectified and insulted, or seen in terms of physical appearance, since they are hardly ever given credit for their intelligence and accomplishments, but I will analyse and focus on this aspect in the final part of my work. As regards the collocates of this search, which can be observed in Figure 13, 'beaut*' is mainly used in monologues and dialogues about beauty in general and what the respective characters perceive as beautiful things.

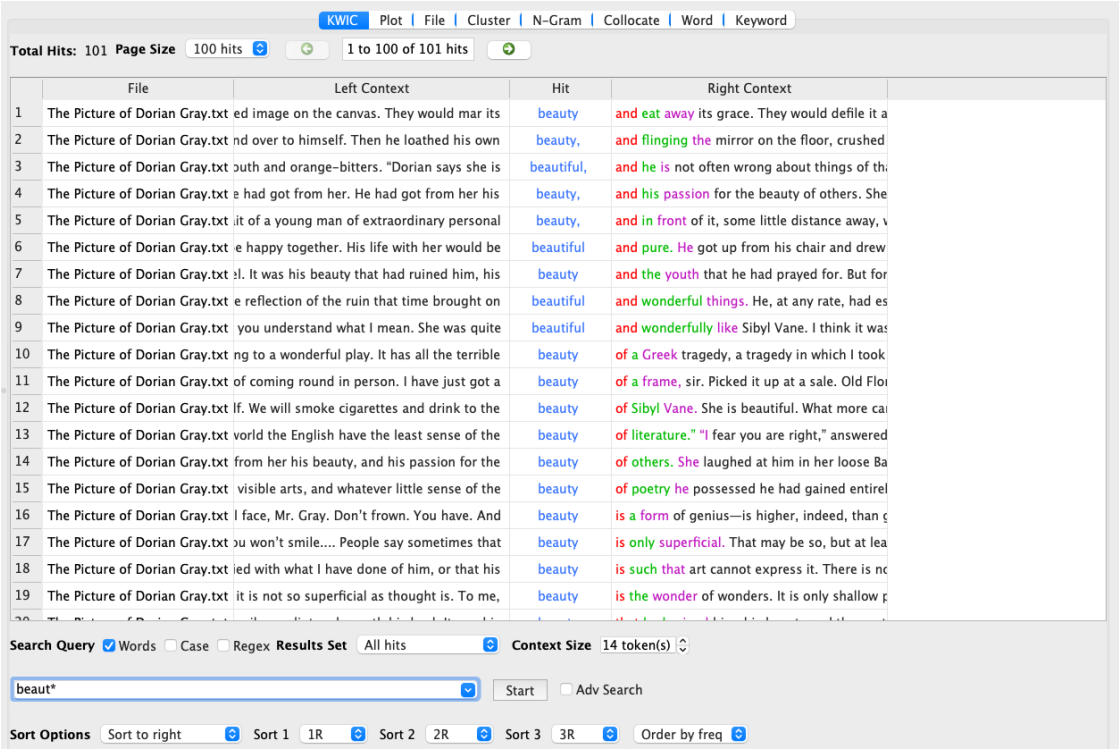


Figure 12. Hits for the word 'beaut*' in *The Picture of Dorian Gray*.

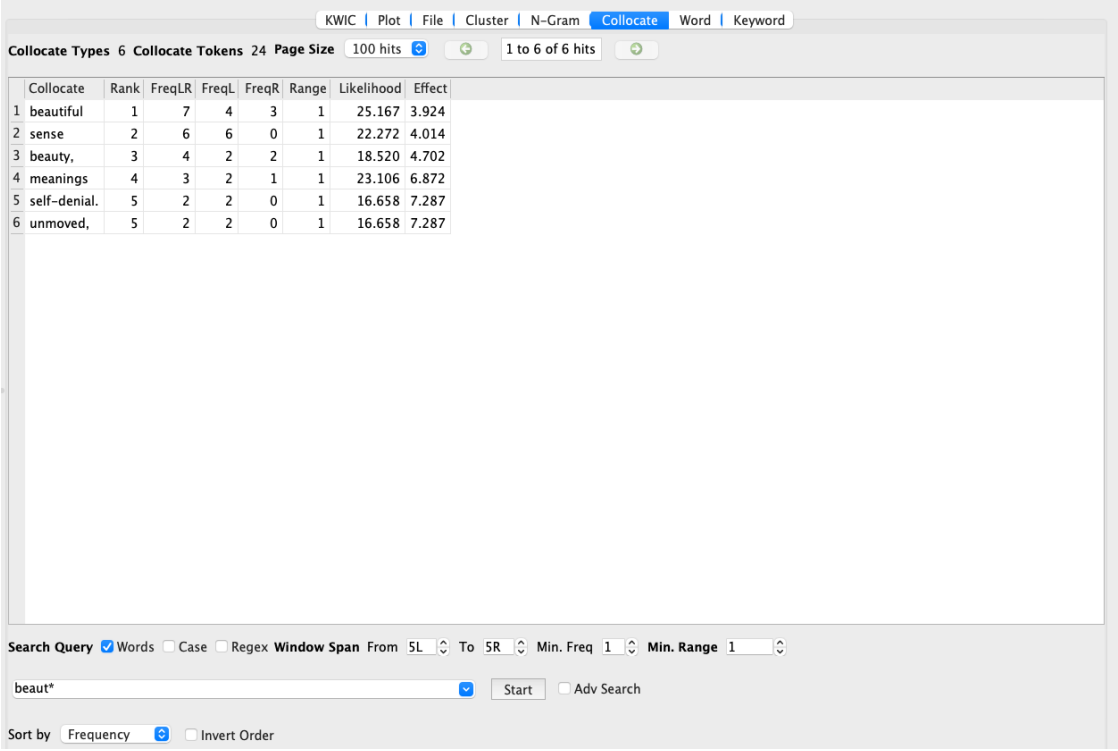


Figure 13. Collocates of the word 'beaut*' in *The Picture of Dorian Gray*.

Subsequently, I repeated the same analysis in *Wilde*, and it revealed a considerable amount of hits (384) throughout the texts that are included in the reference corpus, as Figure 14 shows. However, as with the analysis carried out in the target corpus, the search carried out in the reference corpus did not lead to the discovery and observation of many comments against women, but it mainly showed statements on beauty more at large and beautiful things. This is even demonstrated by the collocate analysis (see Figure 15).

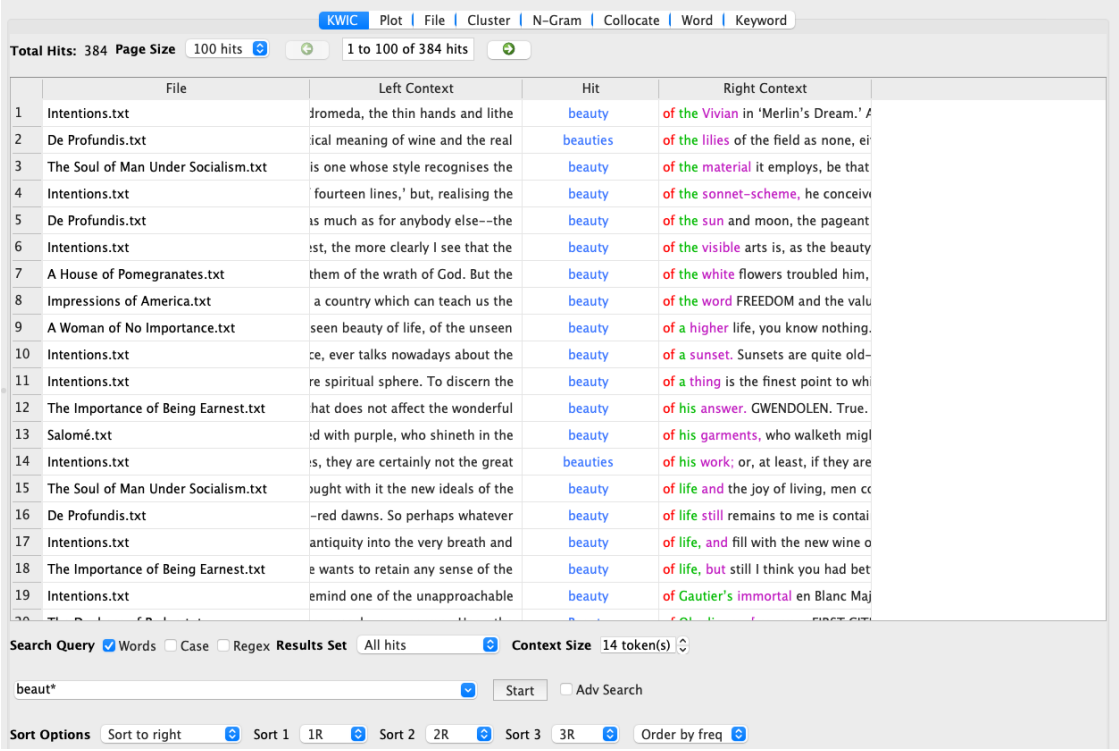


Figure 14. Hits for the word 'beaut*' in *Wilde*.

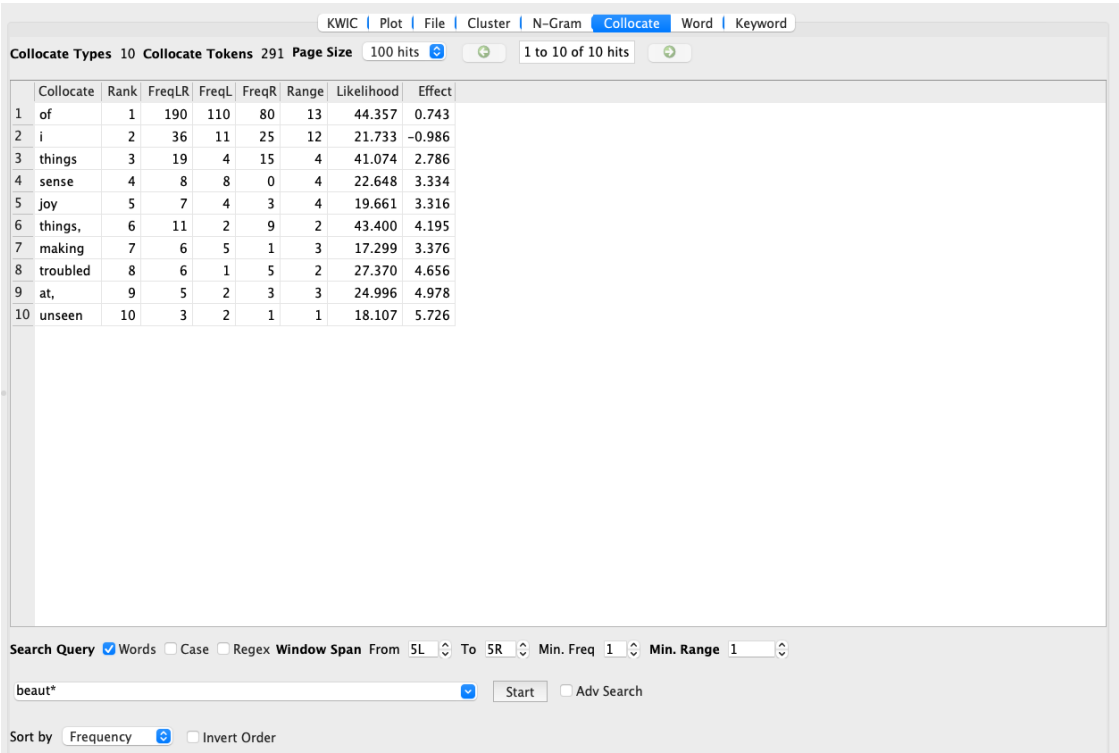


Figure 15. Collocates of the word 'beaut*' in *Wilde*.

The final step of my analysis consists in searching for words denoting sharpness of mind, because I wanted to investigate whether a woman's intellect is celebrated as much as a man's one in *The Picture of Dorian Gray*. Personally, as previously stated, I do not expect to find many examples of women being praised for their intelligence in Wilde's novel. I decided to search for the word 'clever', and my expectations were confirmed, as shown in Figure 16, because this adjective is used in a positive way mainly when it refers to men (see lines 1 and 8 in Figure 16). On the other hand, if women are described with it, it has a negative connotation. For instance, women "are clever at concealing" (see line 2 in Figure 16), it is surprising when women are "too clever" (see line 4), and their "clever tongue" (see line 6) irritates others.

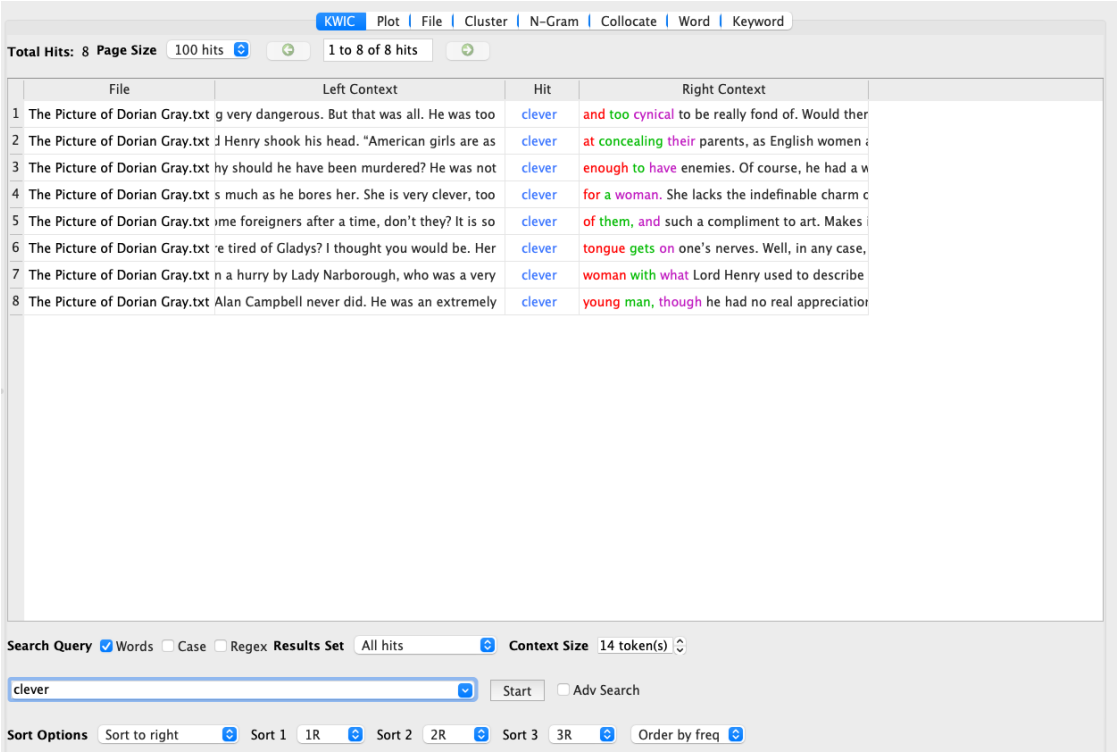


Figure 16. Hits for the word 'clever' in *The Picture of Dorian Gray*.

On the other hand, as it can be observed in the results illustrated in Figure 17, in the reference corpus some hits show how the adjective 'clever' is used for women in an appreciative way (for instance, see lines 3 and 5). However, in other cases the same adjective referred to women is also used in a negative way or context. For instance, in

line 7, ‘clever’ co-occurs with the adverb ‘awfully’, and line 14 says that “women are wonderfully clever in concealing their parents”, which is not exactly a compliment.
In addition, some of the utterances are delivered by other female characters, and this detail is important because there are not many men who celebrate women and their accomplishments. Moreover, this might even suggest that in other literary works of Oscar Wilde’s, women are given more relevance, meaning that they might play a bigger role, while in *The Picture of Dorian Gray* there are no prominent female characters.

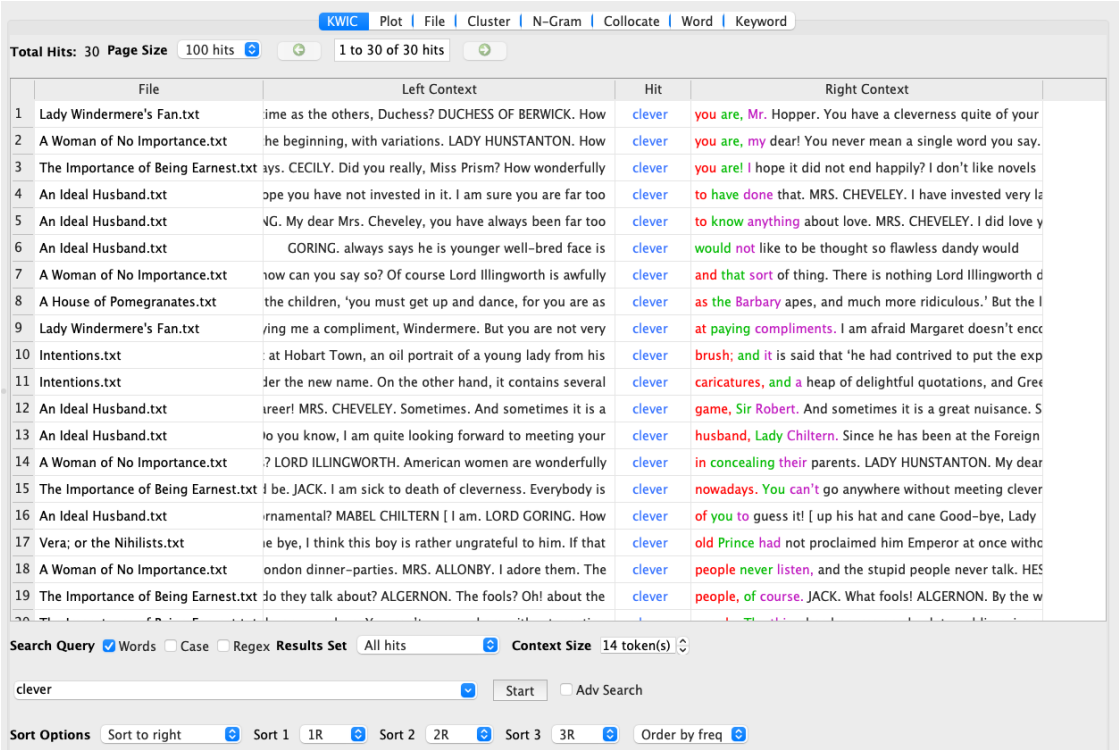


Figure 17. Hits for the word ‘clever’ in *Wilde*.

After having searched various terms in both the target corpus and the reference corpus, *Wilde*, using *AntConc*, I believe that I collected enough results to conclude my corpus-assisted analysis. Hence, in the following part of my thesis, I will assess and evaluate the results obtained in the research, also reflecting on my initial expectation.

3.3. Comparing and contrasting expectations with findings

In the final section of my dissertation, I reflect on my work, comparing and contrasting my initial, *a priori* expectations with the actual outcomes of the research that I conducted. As a matter of fact, summing up the results of my corpus-assisted analysis, some of my initial assumptions were confirmed by the findings, due to the fact that through the observation of concordances and collocates, I realized that *The Picture of Dorian Gray* has several misogynistic aspects and comments. Nonetheless, other findings disclosed surprising information as regards both the target corpus and *Wilde*.

Among the results that surprised me the most, I did not expect to discover that, in *The Picture of Dorian Gray*, the words 'charm' and 'charming' are hardly ever used for a woman in a positive context. What I mean to say is that I knew that women are not often praised in this literary text, but I expected to find more results involving this specific noun or this specific adjective employed with a negation. However, I was pleasantly surprised to find out that 'charm*' has a much higher frequency in *The Picture of Dorian Gray* than in the reference corpora. At the same time, I also wanted to highlight that I was particularly impressed by a result that emerged from the search for 'charm*' in *Wilde*, and that is Mrs Cheveley's utterance in *An Ideal Husband* (1893), which was already reported in this chapter. As I stated in the previous section of this chapter, in this line a female character is trying to use her voice to stand up against the comments of another character, who is a man and is trying to insult and degrade her. Although Mrs Cheveley's character can be described as manipulative, on one hand, it is interesting to see her trying to speak up for herself and, more generally, to read about a female character who has a more active role in the play, but on the other hand, it still depicts women in a negative way, that is as calculating and devious.

In addition, other unexpected results emerged from the search for 'beaut*' in Wilde's novel, because even though my personal expectations as regards the total number of hits were confirmed, I was convinced that even more of them would refer to women. Therefore, the results of this specific search for the collocations in the target corpus were below my expectations, in view of the fact that most of the results of the analysis displayed comments on beauty and beautiful things in broad terms.

Lastly, my assumptions about how women intelligence is portrayed in the target corpus were supported by the results. Indeed, searching the adjective 'clever' did not reveal a copious amount of hits in *The Picture of Dorian Gray*, since there were 8 hits, and in none of them were women and their intelligence celebrated. This term was used in a positive way and context only if referred to a male character, whereas when referred to women its meaning was close to scheming and Machiavellian, similarly to Mrs Cheveley's case in *An Ideal Husband* (1893).

Women are mentioned a high number of times in *The Picture of Dorian Gray*, which is quite interesting, since female characters barely have lines, and they are not the heroines in the narration, they are underrepresented. In this regard, women really seem to be, similarly to Wilde's words, a decoration, a decorative sex. Indeed, as emerged from the various steps of my analysis, the female characters of the novel are often disregarded and not respected. Therefore, *The Picture of Dorian Gray* is a misogynistic novel that is supposed to mirror Victorian England's sexism, especially through Lord Henry's remarks. Nonetheless, it is important to remember that authors and the characters they create in their works are not interchangeable. In other words, characters are tools that authors use to send a message, and in this case, Wilde uses Lord Henry to criticise and expose Victorian society.

Conclusion

Wilde's *The Picture of Dorian Gray* appears to have several misogynistic aspects. My corpus-assisted analysis confirmed that women are often mentioned in the novel, but they rarely speak or play a crucial role. However, when female characters are the main topic of discussion, they are objectified and degraded, and, as pointed out by Kersten (2014: 27), Lord Henry is "especially verbal on this matter", since he is the male character who speaks badly about women the most. My research showed that women are hardly ever praised or given credit neither for their appearance and physical beauty, nor for their intelligence and talents. This latter aspect is even more rare, considering that in the novel it is almost completely absent, as demonstrated by my analysis. As a matter of fact, the search for 'clever' showed that this adjective has a positive connotation mainly when it refers to men, while it is used negatively to describe women.

This kind of mistreatment and inequality mirrors the actual position of women in Victorian society. Indeed, Mohammed and Abdullah (2019: 537) state that "the novel is a reflection of the long and deep-rooted sexism in the English language and gives a clear image of the position and the status of women in the Victorian age". However, as previously discussed in this thesis, in other works of Wilde's, for example in *An Ideal Husband* (1893), there are some cases (but not many) in which women try to stand up for themselves. This can be considered as an indicator of the fact that the novel *The Picture of Dorian Gray* could be Wilde's most misogynistic work, or at least it seems to address misogyny through its characters in a more direct way.

While countless of literary critics and experts have examined Wilde's prose, not many stylisticians have analysed *The Picture of Dorian Gray* with the help of corpus techniques. This is a pity considering the potential aspects and themes of the novel that could be observed, but at the same time it is understandable and logical, because it is important to remember that corpus stylistics is still a fairly young discipline. For further research, in the future, it would be interesting to conduct similar corpus stylistic analyses to observe masculinity in this literary work, the theme of duplicity, the conflict between youth and old age, as well as the one between eternity and life, the mortality of beauty, or Lord Henry's rhetoric and influence.

These are only some instances, because the versatility of corpus stylistics allows to explore innumerable aspects or features of a literary work. My dissertation itself is only an example, and it is one specific to Wilde and his novel, and it provides some insights into the much larger set of functions that are typical of corpus stylistics. Reflecting on my work, deciding to involve software programmes was extremely useful for the goal of my investigation. Indeed, corpus techniques allowed me to examine the whole text of *The Picture of Dorian Gray* instead of small extracts, answering my research questions. The fact that the corpora were stored electronically is what made it possible to analyse the novel systematically, with the aid of computer tools. This was a huge advantage, because I could observe the frequency of the words that I intended to analyse, how they are distributed throughout the entire text, and how they co-occur with other words.

Bibliography

Anthony, L. (2022). AntConc (Version 4.0.3) [Macintosh]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/ (accessed 5 February 2022).

Beckson, K. (2021), Oscar Wilde. *Encyclopedia Britannica*. https://www.britannica.com/biography/Oscar-Wilde (accessed 15 January 2022)

Biber, D. (2012), Corpus-Based and Corpus-driven Analyses of Language Variation and Use. Oxford Handbooks Online.

Biber, D. (2011), Corpus Linguistics and the study of Literature: back to the future?. In *The Scientific Study of Literature*, v. 1, n. 11, p. 15-23.

Biber, D.; Johansson, S.; Leech, G.; Conrad, S.; & Finegan, E. (1999), Longman Grammar of Spoken and Written English. Harlow: Pearson Education.

Binongo, J. N. G.; Smith, M. W. A. (1999), A bridge between statistics and literature: the graphs of Oscar Wilde's literary genres.

Buzwell, G. (2014), Daughters of decadence: the New Woman in the Victorian fin de siècle. *British Library*. https://www.bl.uk/romantics-and-victorians/articles/daughters-of-decadence-the-new-woman-in-the-victorian-fin-de-siecle (accessed 29 January 2022).

Cheng, W. (2013), Semantic Prosody. In C. A. Chapelle (ed.) *The encyclopedia of applied linguistics*. Oxford, England: Wiley-Blackwell Publishing Ltd., 5178-5184.

Culpeper, J. (2002), Computers, language and characterization: An analysis of six characters in *Romeo and Juliet*. In U. Melander-Marttala, C. Ötsman and M. Kytö (eds), *Conversation in Life and in Literature*, p. 11-30. Uppsala: Universitetstryckeriet.

Cvrček, V.; Fidler, M. (2019), Chapter 5. More than keywords: Discourse prominence analysis of the Russian Web Portal *Sputnik Czech Republic*. In M. Berrocal and A. Salamurović (eds.) *Political Discourse in Central, Eastern and Balkan Europe*, p. 93-117. Amsterdam/Philadelphia: John Benjamins.

Evans, D. (2018), Corpus building and investigation for the Humanities.

Fischer-Starcke, B. (2010), Corpus Linguistics in Literary Analysis: Jane Austen and her Contemporaries. London: Continuum.

Grinstein, A. (1980), Oscar Wilde. *American Imago*, v. 37, n. 2, p. 125-179. http://www.jstor.org/stable/26303564 (accessed 15 January 2022).

Hussein, K. S.; Abdul-Ameer A. H. (2017), A Corpus-Driven Approach to Stylistic Analysis of a Lexical Richness Curve.

Junmei, J. (2017), A Corpus-Aided Study of Language Features of 'The Importance of Being Earnest'. https://files.eric.ed.gov/fulltext/EJ1269091.pdf (accessed 17 January).

Kersten, A. (2014), Gender and Sexuality in *The Picture of Dorian Gray* and *Dorian Gray*. Bachelor thesis, Radboud University, Nimega. https://theses.ubn.ru.nl/bitstream/handle/123456789/809/Kersten%2C_A.M.C._1.pdf?sequence=2 (accessed 8 December 2021).

Mahlberg, M.; Smith, C. (2010), Corpus approaches to prose fiction: Civility and body language in *Pride and Prejudice*. In D. McIntyre & B. Busse (Eds.), *Language and Style*, p. 449-167. Basingstoke: Palgrave Macmillan.

Mahlberg, M. (2010), Corpus Linguistics and the Study of Nineteenth-Century Fiction, Journal of Victorian Culture, v. 15, n. 2, p. 292-298.

Mahlberg, M. (2012). Corpus Stylistics and Dickens's Fiction (1st ed.). New York and London: Routledge.

Mahlberg, M. (2007b), Corpus stylistics: bridging the gap between linguistic and literary studies. In M. Hoey, M. Mahlberg, M. Stubbs & W. Teubert (Eds.), *Text, Discourse and Corpora*, p. 191-218. London: Continuum.

McEnery, T., Xiao, R., & Tono, Y. (2006). Corpus-Based Language Studies - An Advanced Resource Book. New York and London: Routledge.

Mills, S. (1995), Feminist Stylistics. London: Routledge.

Mohammed, M. B.; Abdullah, A. A. (2019), Sexism in Wilde's *The Picture of Dorian Gray*: Linguistic Analysis. Journal of Tikrit University for Humanities, v. 7, n. 2, p. 531-538. https://journals.indexcopernicus.com/api/file/viewByFileId/941412.pdf (accessed 4 January 2022).

Nordentoft, S. J. (2020), Dorian Gray and the Wilde Senses: A corpus Stylistic Analysis of Oscar Wilde's *The Picture of Dorian Gray*. Master thesis, Aalborg University, Aalborg.
https://projekter.aau.dk/projekter/files/334899360/Dorian_Gray_and_the_Wilde_Senses_Signe_Nordentoft1.pdf (accessed 18 January 2022).

Parfitt, G. (2013), The Picture of Dorian Gray Plot Summary. https://www.litcharts.com/lit/the-picture-of-dorian-gray/summary (accessed 21 January 2022).

Parker, C. (1995), Gender Roles and Sexuality in Victorian Literature. Aldershot: Scholar.

Pearson, J. (1998), Terms in Context. Amsterdam: John Benjamins.

Scott, M.; Tribble, C. (2006), Textual Patterns. Key Words and Corpus Analysis in Language Education. Amsterdam: John Benjamins.

Sinclair, J. (1987), Looking up: An Account of the COBUILD projects in lexical computing and the development of the Collins COBUILD English Language Dictionary. London, England: Collins.

Tobin, G. B. (2019), Oscar Wilde Biography. https://www.wilde-online.info/oscar-wilde-biography.htm (accessed 4 January 2022).

Tognini-Bonelli, E. (2001), Corpus Linguistics at Work. Amsterdam: John Benjamins.

Wang, S. (2016), Misogyny and Female Stereotyping in *The Picture of Dorian Gray* and *Venus in Furs*.

Wilde, O. (1891), *The Picture of Dorian Gray*. Ed. Alma Classics Ltd (2014).

Summary in Italian

La mia tesi propone un'analisi degli aspetti misogini presenti nell'opera *Il ritratto di Dorian Gray* (1891) di Oscar Wilde (1854-1900). Più precisamente, ho deciso di condurre la mia indagine basandomi sui principi e sulla metodologia di una disciplina emergente che in inglese è conosciuta come *corpus stylistics*. Tale insegnamento nasce dall'unione di altre due discipline: *stylistics* e *corpus linguistics*. La prima si occupa dello studio linguistico dello stile di testi letterari, mentre la seconda è una disciplina linguistica che comporta l'utilizzo di software per l'analisi in questione. Non a caso, questo tipo di ricerche si basa sull'analisi di un *corpus*, ovvero una raccolta di testi creata con attenzione, per uno scopo ben preciso e conservata in formato elettronico. Tuttavia, per affermare che una determinata caratteristica ha un certo valore o una certa sfumatura di significato all'interno del testo che si vuole analizzare, è necessario avere due *corpora*. Uno di questi, il *target corpus*, comprende il testo che si intende analizzare, mentre il secondo, il *reference corpus*, è più generale, comprende un maggior numero di testi, in modo tale che questo rappresenti un determinato tipo di linguaggio nel modo più completo possibile. Dunque, con '*corpus stylistics*' si intende l'utilizzo di *corpora*, degli strumenti e dei metodi tipici della disciplina conosciuta come *corpus linguistics*, applicato allo studio di testi letterari. Di fatto, uno dei vantaggi principali degli studi di *corpus stylistics* è che, rispetto ad altre discipline, come la linguistica e gli studi letterari, permettono l'analisi elettronica, sistematica di intere opere letterarie e non solo di estratti.

Nel mio elaborato riporto quali sono le metodologie e gli approcci più frequenti di questa disciplina, soffermandomi soprattutto sulla differenza tra un'analisi *corpus-assisted* ed una *corpus-driven*. Con '*corpus-assisted*' si intende una ricerca che ha come scopo principale la validazione, la confutazione o il perfezionamento di teorie, ipotesi ed intuizioni preliminari. Invece, la terminologia '*corpus-driven*' indica un approccio che esplora i *corpora* in modo più induttivo, scoprendo nuove costruzioni linguistiche senza avere presupposizioni. I principali software che vengono usati in molte analisi di *corpus stylistics* sono *AntConc*, *kfNgram*, *Wmatrix* e *WordSmith Tools*.
La maggior parte delle ricerche riconducibili a quest'ambito di studio ha come scopo principale l'osservazione e l'analisi delle ripetizioni e della distribuzione delle parole che

sono caratteristiche di un autore, di un testo, o anche di un singolo personaggio. I tre approcci metodologici più utilizzati per tali ricerche sono l'analisi di *keywords*, l'identificazione di *extended lexical phrases* e l'analisi di *collocations*. L'analisi di *keywords* evidenzia quali sono le singole parole distintive di un'opera letteraria, mentre nel caso di un'analisi di *extended lexical phrases* vengono osservate le sequenze lessicali che prevalgono in un testo. Invece, l'analisi di *collocations* si concentra sulle parole che tendono a comparire insieme più frequentemente di quanto ci si potrebbe aspettare.

Nella mia tesi vengono presentati l'autore Oscar Wilde e la sua opera *Il ritratto di Dorian Gray*, che contiene numerosi esempi di misoginia, stereotipi sulle donne e sessismo. Ciò nonostante, è doveroso ricordare che questi commenti pronunciati dai personaggi del romanzo intendono rispecchiare la società vittoriana, ovvero quella società in cui lo stesso Wilde viveva e che voleva smascherare tramite le sue opere.
Per l'analisi dell'opera, ho creato due *corpora*: il *target corpus* è formato dal testo de *Il ritratto di Dorian Gray*, mentre il *reference corpus*, chiamato *Wilde*, comprende 15 testi di cui Wilde è l'autore. Per quanto riguarda la scelta del software, ho deciso di utilizzare *AntConc*, un programma in grado di esaminare i *corpora*, gli elementi che li compongono, la frequenza delle varie unità linguistiche e la loro distribuzione. Inoltre, è necessario precisare che, con l'aiuto dei *corpora* e di *AntConc*, per la mia analisi intendo optare per un approccio *corpus-assisted*, dal momento che voglio testare la validità delle mie intuizioni personali. Presumo che parole come '*woman*' e '*women*' compaiano numerose volte nel romanzo, soprattutto nei discorsi di Dorian Gray e Lord Henry. Per di più, leggendo l'opera per la prima volta, ho avuto l'impressione che una buona parte di queste conversazioni avessero come argomento la bellezza esteriore e l'apparenza delle donne, ma anche lo charm ed il fascino. Dunque, mi aspetto che siano numerosi gli aggettivi che descrivono queste qualità, magari anche con dei negativi, come '*not*', in contesti dispregiativi o con accezioni negative. In generale, mi aspetto di osservare diverse parole ed espressioni che mirano all'oggettificazione delle donne, viste come inferiori, meno capaci o intelligenti rispetto agli uomini.

L'inizio della mia analisi consiste nella ricerca dei termini '*woman*' e '*women*', che in effetti si sono rivelati piuttosto frequenti all'interno de *Il ritratto di Dorian Gray*. I

risultati mostrano anche che in molti di questi è presente un contesto negativo, dal momento che, ad esempio, le donne vengono giudicate per il loro aspetto fisico o età e non vengono apprezzate per la loro intelligenza. La medesima ricerca condotta nel *reference corpus*, *Wilde*, dimostra che anche in questo caso le due parole in questione sono piuttosto frequenti e che spesso si tratta di commenti misogini.

La mia tesi procede con la ricerca del nome '*Sibyl*', una delle poche figure femminili presenti nel romanzo di Wilde, e per questo motivo ero curiosa di osservare quali sono le parole più utilizzate con il suo nome. Ad ogni modo, non mi aspetto di trovare il nome '*Sibyl*' in *Wilde*, siccome nessun altro dei suoi personaggi ha lo stesso nome. I risultati hanno mostrato che, ne *Il ritratto di Dorian Gray*, il nome '*Sibyl*' viene associato ed affiancato in più di un'occasione alla morte e al suicidio, ma in alcuni casi anche all'amore e al matrimonio. Inoltre, è possibile notare anche un paio di complimenti sulle sue capacità di recitazione, ma questi casi sono limitati, poiché Dorian Gray inizialmente la reputa talentuosa, ma in un secondo momento cambia drasticamente idea.

Successivamente, segue la ricerca della parola '*charm**', che mi aveva dato l'impressione di essere molto ricorrente nel testo, ed il simbolo ' *' è il cosiddetto '*wildcart character*', ovvero un simbolo che sta per uno o più caratteri che seguono le lettere precedenti. Questa funzione è estremamente utile, perché, tramite un'unica ricerca, permette di osservare i risultati delle varie forme e coniugazioni della parola di base, come '*charming*', '*charmed*', '*charms*', etc. In effetti, dalla ricerca emerge che questo termine, nelle sue molteplici forme e coniugazioni, compare ripetutamente nel *target corpus*, ma principalmente per complimentare i personaggi maschili o per descrivere in maniera positiva qualcosa che essi trovano piacevole ed interessante. Invece, quando questi termini vengono rivolti a donne, hanno un'accezione negativa. Ripetendo gli stessi passaggi nel *reference corpus*, anche in questo caso ci sono numerosi risultati per quanto riguarda la frequenza di '*charm**', ma comunque nettamente inferiori rispetto al *target corpus*. Oltre a questo, in *Wilde* solo un numero esiguo di risultati si riferisce alle donne, ma c'è uno di questi che, diversamente rispetto agli altr, mette in risalto una frase in cui un personaggio femminile impiega '*charming*' nel tentativo di difendersi dai commenti di un uomo.

In seguito, conduco un'altra ricerca con un *wildcard character*: '*beaut**'. Nel *target corpus*, molti di questi risultati sono associati all'aspetto fisico di Dorian Gray e a discorsi

sul concetto di bellezza in sé, ma solo una parte di questi si riferisce alle donne in maniera positiva. Questo aspetto viene confermato anche dall'analisi di *collocations*. Quindi, è possibile notare che sono i personaggi maschili ad avere più rilevanza in questo testo letterario. In aggiunta, quando le donne vengono menzionate, nella maggior parte dei casi non si tratta di apprezzamenti nei loro confronti, in quanto vengono per lo più denigrate e se vengono notate positivamente, è quasi sempre per il loro aspetto fisico e non per la loro personalità, intelligenza o abilità. La ricerca di '*beaut**' in *Wilde*, nonostante i numerosi risultati, mostra che anche in questo caso si tratta principalmente di discorsi sulla bellezza in sé. Come per la medesima ricerca nel *target corpus*, pure in *Wilde* l'analisi di *collocations* supporta questa argomentazione.
Infine, ho cercato nei due *corpora* la parola '*clever*' per osservare se all'intelligenza delle donne viene data tanta importanza quanto a quella degli uomini. Ne *Il ritratto di Dorian Gray*, questo aggettivo viene impiegato principalmente per descrivere i personaggi maschili in maniera positiva, mentre se riferito a quelli femminili ha decisamente una connotazione negativa. Si ottengono risultati simili nel *reference corpus*, anche se compaiono un paio di esempi in cui l'aggettivo viene usato per complimentare le donne.

Dopo aver osservato e riflettuto sui risultati che sono emersi dai vari passaggi della mia analisi, paragonandoli alle mie aspettative iniziali, è possibile notare che in alcuni casi le mie ipotesi sono state confermate, mentre in altri ci sono stati dei risultati inaspettati. Infatti, pur essendo a conoscenza del fatto che le donne non vengono celebrate spesso ne *Il ritratto di Dorian Gray*, mi aspettavo di trovare ancora più commenti negativi nei confronti delle donne con il termine '*charm**'. Tale ricerca mi ha sorpreso anche per un altro aspetto, ovvero non immaginavo che '*charm**' fosse piuttosto frequente anche nel *reference corpus* e non mi aspettavo di trovare addirittura un esempio in cui un personaggio femminile utilizza questa parola per difendersi dai commenti di un uomo. Per di più, la ricerca '*beaut**' ha inaspettatamente portato all'osservazione del fatto che, nel *target corpus*, tale termine viene effettivamente impiegato ripetutamente, come avevo previsto, ma si tratta per lo più di conversazioni sul concetto di bellezza in generale.
Tuttavia, le parole analizzate sono effettivamente risultate piuttosto comuni e centrali all'interno de *Il ritratto di Dorian Gray*. Le donne ed i personaggi femminili, come avevo

inizialmente previsto, vengono menzionati frequentemente nel corso dell'opera e nella maggior parte dei casi si tratta di commenti misogini, sessisti ed irrispettosi.