



Universiteit  
Leiden  
The Netherlands

# Reading Jane Austen in the Digital Age: Exploring the Implications of Textual Analysis Tools for Literary Scholarship

Veenen, Myrthe Van

## Citation

Veenen, M. V. (2024). *Reading Jane Austen in the Digital Age: Exploring the Implications of Textual Analysis Tools for Literary Scholarship*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master Thesis, 2023](#)

Downloaded from: <https://hdl.handle.net/1887/3775173>

**Note:** To cite this publication please use the final published version (if applicable).

**Reading Jane Austen in the Digital Age:  
Exploring the Implications of Textual Analysis Tools for Literary Scholarship**



MA Thesis – Book and Digital Media Studies  
Myrthe van Veenen  
s3496546

Dr. Peter Verhaar  
Dr. Fleur Praal  
Wordcount: 21.915  
Date: 14 June 2024

### Abstract

‘This thesis explores the transformative impact of digital mediums and computational textual analysis tools on the reading and interpretation of Jane Austen's works. Through a comprehensive examination, the study contrasts traditional print reading with digital methods, emphasizing the shift in the literary landscape. Chapter 1 introduces the research questions and relevant theoretical framework, focusing on how digital tools influence the comprehension and appreciation of Austen's texts. Chapter 2 delves into the historical context, examining the reception and interpretation of Austen's novels by established critics through close reading methods. Chapter 3 presents a detailed case study, utilising digital textual analysis to uncover patterns and themes in Austen's writing. The conclusion synthesises these findings, highlighting the enhanced efficiency of textual analysis, and the unveiling of new aspects in Austen's texts. The research demonstrates the value of digital tools in providing empirical evidence and new insights, enriching the field of literary criticism. It also addresses the relevance and future potential of computational methods, proposing directions for further research and acknowledging the limitations of computational research in the literary field and the current study. This work underscores the significant role of digital innovation in reshaping the study and appreciation of classic literature.’

**INDEX**

<b>CHAPTER 1 – INTRODUCTION .....</b>	<b>4</b>
1.1. <i>INTRODUCTION</i> .....	4
1.2. <i>COMPUTATIONAL LITERARY STUDIES</i> .....	5
1.3. <i>RESEARCH QUESTIONS</i> .....	6
1.4. <i>DISTANT READING VS CLOSE READING</i> .....	7
1.5. <i>INFLUENCES OF MEDIUM</i> .....	10
1.6. <i>STRUCTURE OF THE THESIS</i> .....	12
<b>CHAPTER 2 – READING JANE AUSTEN IN THE PRINT AGE.....</b>	<b>15</b>
2.1. <i>INTRODUCTION</i> .....	15
2.2. <i>VIEWS ON AUSTEN’S POSITION IN THE CANON OF ENGLISH LITERATURE</i> .....	16
2.3. <i>LITERARY GENRES</i> .....	18
2.4. <i>THEMATIC CONCERNS</i> .....	20
2.5. <i>CHARACTERISATION</i> .....	22
2.6. <i>DICTION</i> .....	25
2.7. <i>LITERARY DEVICES</i> .....	27
2.8. <i>RECEPTION OF AUSTEN’S WORK FROM THE 19<sup>TH</sup> CENTURY ONWARDS</i> .....	30
<b>CHAPTER 3 – READING JANE AUSTEN IN THE DIGITAL AGE.....</b>	<b>35</b>
3.1. <i>INTRODUCTION</i> .....	35
3.2. <i>COMPUTATIONAL CRITICISM ON AUSTEN</i> .....	36
3.3. <i>METHODOLOGY</i> .....	39
3.4. <i>DICTION AND VOCABULARY</i> .....	41
3.5. <i>LINGUISTIC STRUCTURE</i> .....	44
3.6. <i>FIGURES OF SPEECH</i> .....	50
3.7. <i>SEMANTIC CATEGORIES</i> .....	52
3.8. <i>LITERARY THEMES</i> .....	56
3.9. <i>CONCLUSION</i> .....	59
<b>CHAPTER 4 – CONCLUSION .....</b>	<b>61</b>
4.1. <i>INTRODUCTION</i> .....	61
4.2. <i>DISCUSSION</i> .....	61
4.3. <i>LIMITATIONS AND FUTURE RESEARCH</i> .....	66
<b>BIBLIOGRAPHY .....</b>	<b>69</b>
<b>APPENDICES .....</b>	<b>72</b>
APPENDIX A .....	72
APPENDIX B .....	76

## Chapter 1 – Introduction

### 1.1. Introduction

It is a truth universally acknowledged that the advent of the digital age has ushered in a multitude of new possibilities and computational textual analysis tools that are transforming the way in which we interact with information and literature. As technology continues to shape our reading practices, it becomes imperative to examine the implications of these advancements on our experience of literature, particularly in the context of renowned authors such as Jane Austen, Charles Dickens, and William Shakespeare. This thesis aims to explore the effects of the digital medium and its associated textual analysis tools on the act of reading, and by extension literary criticism. These digital tools, and the associated new form of humanities, allow for researchers to approach familiar texts and actions, e.g., literary analysis, in a new way. Most of the research that has been done regarding literary devices rests on close reading analysis and methods.<sup>1</sup> In the field of humanities and literary studies, however, the use of computational methods for literary analysis has grown significantly.<sup>2</sup> This thesis will explore the changing landscape of reading, literary criticism and the ways in which technology shapes our engagement with literary texts. This thesis will base itself predominantly on academic works which will aid in the understanding of computational research in literary studies. Additionally, other sources will be used to create a better understanding of Jane Austen as an author and traditional literary criticism surrounding her writing. These sources put together encompass different approaches and results which could be tested against in the case study of this project.

By experimenting with the computational detection and comparing findings to results gathered through traditional methods such as close reading, this study can enhance our comprehension of literary texts, uncover hidden layers of meaning, and shed light on the intricate interplay between language, culture, and human experience as experienced by the characters in Austen's novels. This thesis mainly investigates how textual analysis tools affect the reading experience for both the general reader and the literary scholar by using Jane Austen as a case study. As an established author in the English literature canon, she provides a stable basis to test the theory on, using her novels as primary sources, namely, *Pride and Prejudice*, *Sense and Sensibility*, *Emma*, *Northanger Abbey*, *Mansfield Park*, and *Persuasion*.

---

<sup>1</sup> P. Verhaar, "Affordances and Limitations of Algorithmic Criticism", (unpublished doctoral thesis, 2016), p. 247.

<sup>2</sup> *ibid.*, p. 22.

### *1.2. Computational Literary Studies*

In an age marked by technological leaps and digital proliferation, the realm of literary analysis has expanded beyond the traditional confines of close reading and manual interpretation. A new area has emerged with the integration of computational theory into the study of literature, fundamentally altering the environment of literary scholarship. Computational theory, rooted in the principles of computer science and mathematical modelling offers a novel lens through which to perceive and interpret the intricacies of literary works.<sup>3</sup> As Martin Paul Eve explains, the ‘methods are claimed to permit us, at a distance, to ingest, process, and perhaps understand texts within grand perspectives.’<sup>4</sup> By harnessing the analytical capabilities of computers, researchers are empowered to navigate the vast textual field, identify patterns, and discern semantic nuances that might elude traditional modes of analysis. This is accomplished by computational algorithms that are designed to recognize patterns, frequencies, and relationships within a text by analysing the words that are used. Elements such as word frequencies, syntactic structures, sentiment analysis, and stylistic patterns become amenable to rigorous measurement and analysis. By quantifying these features, researchers can discern shifts in an author’s writing style over time, identify intertextual references, and even attribute authorship to disputed works.<sup>5</sup> As Peter Verhaar explains, once ‘individual textual units have been identified, they can also be analysed quantitatively’.<sup>6</sup> This all contributes to a deeper understanding of the literary landscape. However, the technological terrain in literary studies is still a developing one. As such, algorithms for textual analysis are still limited in more complicated areas, such as recognizing literary devices as irony, or metaphors. Because to be able to ‘quantify a phenomenon, it is important, firstly, to ensure that the computer can recognize instances of this phenomenon’.<sup>7</sup> John Burrows notes that, in computational analysis, ‘statistical comparisons are meaningful only to the extent that they rest upon an appropriate common base’.<sup>8</sup> This means that it is necessary to establish such a common base in a research project, such as a specific author or time period. It remains difficult to incorporate cultural, or gender nuances to which the human scholar is sensitive, into a computational tool or algorithm.

---

<sup>3</sup> M. P. Eve, *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell’s Cloud Atlas*, (Stanford: Stanford University Press, 2019), p. 2.

<sup>4</sup> *ibid.*

<sup>5</sup> Verhaar, p. 31.

<sup>6</sup> *ibid.*, p. 104.

<sup>7</sup> *ibid.*, p. 109.

<sup>8</sup> J.F. Burrows, *Computation into Criticism: A Study of Jane Austen’s Novels and an Experiment in Method* (New York: Oxford University Press, 1987), p. 41.

The integration of computational theory into literary scholarship has ushered in a profound renaissance, reshaping the very essence of how scholars' approach and unravel the intricate tapestries of literary works. While traditional close reading remains an indispensable tool, its meticulous scope often skirts over intricacies which can only be observed when collections of texts are viewed at a distance. Computational analysis, in contrast, stands poised to unearth patterns, repetitions, and intertextual resonances that enrich the texture of the text. The integration of computational theory into literary scholarship does kindle discussions on algorithmic bias, the preservation of digital texts, and the profound implications of applying digital tools to examine the intent of the author. However, the main question is: how much does it also impact the reading experience?

### *1.3. Research Questions.*

In the ever-evolving department of literary scholarship, the digital medium and its associated textual analysis tools have emerged as transformative forces, reshaping the way scholars engage with and interpret literary works. This thesis embarks on a multifaceted exploration of the profound effects of the digital age on the reading experience of the literary critic, specifically focussing on Jane Austen's texts as a lens through which to examine the dynamic interplay between technology and literary criticism, and the nuances of her renowned narratives. Central to this study is the overarching question: *How does the digital medium and the utilisation of textual analysis tools affect the reading experience for literary scholars? What is the relevance and the value of these analysis tools in literary criticism, and what new aspects can be observed or noticed in Jane Austen's texts?*

There are a few central themes which are important to answering the central question to this thesis. Through sub questions, this thesis aims to understand how digital media alter reading Austen's works and how that informs the evolving ways literary scholarship engages with classics. Additionally, in this thesis I will contrast traditional critiques with digital insights to uncover novel perspectives and to challenge prevailing interpretations. Moreover, I hope to show that textual tools can deepen comprehension, unveiling hidden layers and contextual nuances in Austen's narratives. Such new digital insights can enrich the current understanding of Austen's themes, characters, and societal commentaries. However, this thesis also aims to identify tool-related limitations to ensure robust analysis and discernment of potential biases. I hope to prove that digital methods refine close reading skills, enabling nuanced explorations of Austen's intricate literary techniques. If this can be proven, then I can also prove that the impact of digital tools can extend beyond Austen, influencing broader

literary analysis practices and methodologies. In short, this thesis hopes to strike a balance between digital and traditional methods and to ensure comprehensive insights, while avoiding overlooking contextual subtleties.

This thesis aims to contribute to the field of computational literary studies in several ways. It firstly explores the changing field of reading, specifically for literary scholars. This aspect of it is inspired by Prof. Dr. Adriaan van der Weel's research on deep reading and distant reading and the differences between reading on a digital medium versus a print version.<sup>9</sup> Additionally, this research aims to provide a thorough understanding of the integration of computational methods in literary analysis. Informed by Dr Peter Verhaar's study on computational analysis regarding poetry, this research aims to apply similar techniques to novels and investigate whether such methods contribute to literary criticism and if so, how.<sup>10</sup> In doing so, this research builds upon existing methods and research and creates a deeper understanding on the continuously developing region of textual analysis tools, adding the aspect of the reading experience for scholars to this discussion.

#### *1.4. Distant Reading vs Close Reading*

In recent years, literary and cultural analysis has witnessed a transformative shift with the emergence of the concept of "distant reading".<sup>11</sup> Coined by Franco Moretti, distant reading represents a methodological departure from traditional close reading practices, introducing a novel approach to comprehending the intricate tapestry of texts. Distant reading, as Paul Eve notes, 'is concerned with reductive but nonetheless labour-saving methods that use the untiring repeatability of computational tasks to garner statistically informed deductions about novels or other works that one has not read'.<sup>12</sup> Distant reading enables scholars to unveil hidden patterns, overarching trends, and underlying structures that may elude conventional close readings. It involves processing and analysing extensive amounts of text data, often in digital formats, to identify linguistic patterns, genre shifts, historical trends, and other macro-level phenomena.<sup>13</sup> Distant reading can also include the use of visualisations, such as graphs, charts, and maps. These visualisations help researchers identify and communicate patterns

---

<sup>9</sup> A. van der Weel, 'Salient Features of Digital Textuality' in *Changing Our Textual Minds: Towards a Digital Order of Knowledge* (Manchester: Manchester University Press, 2011).

<sup>10</sup> Verhaar, p. 7.

<sup>11</sup> A. van der Weel, 'Feeding Our Reading Machines: From the Typographic Page to the Docuverse', *Digital Studies*, 6 (2017), p. 11.

<sup>12</sup> Eve, p. 3.

<sup>13</sup> *ibid.*

and trends more effectively. Since it is impractical to analyse every text in a large corpus, distant readers often work with representative samples.

However, researchers need to be mindful of potential biases introduced by their choice samples. Distant reading has been applied to various fields beyond literature, including history, linguistics, sociology, and more. It offers a way to explore and understand large-scale cultural and historical changes that may not be immediately evident through traditional qualitative methods. For instance, Julian Brooke, Adam Hammond, and Graeme Hirst researched free indirect discourse in modernist fiction through computational methods.<sup>14</sup> Their research addressed the ‘extent to which social attributes of the various characters (in particular age, class, and gender) are reflected in their lexical stylistic profile’.<sup>15</sup> Another example of a study which employed distant reading methods is that of Mary DeForest and Eric Johnson, who investigated the density of Latinate words in the speeches of Jane Austen’s characters.<sup>16</sup> They propose that Jane Austen deliberately used different densities of Latinate words in the speeches of her characters to form her characters.<sup>17</sup> This is because, as they note, ‘Higher densities of Latinate words sometimes indicate intelligence and moral seriousness, at other times, they expose a character’s formality or hypocrisy. Lower densities indicate lesser intelligence or, in the case of sailors, humble birth.’<sup>18</sup> Both these studies employed computational methods and distant reading to uncover depths to literature on a larger scale than close reading methods would allow.

Traditionally, close reading has been the heart of literary analysis, which became prominent following the rise of New Criticism in the 1920s.<sup>19</sup> Rooted in the meticulous examination of individual texts, close reading offers a profound and intricate understanding of literary works by delving into their nuances, language, and subtext. It is often considered as a cornerstone of literary scholarship, as close reading involves a detailed exploration of the text’s structure, imagery, symbolism, and rhetorical devices.<sup>20</sup> By scrutinising every word, phrase, and punctuation mark, close reading uncovers layers of meaning and invites readers to

---

<sup>14</sup> J. Brooke, A. Hammond, and G. Hirst, ‘Using Models of Lexical Style to Quantify Free Indirect Discourse in Modernist Fiction’, *Digital Scholarship in the Humanities*, 32.2. (2017), pp. 234-50, (p. 234).

<sup>15</sup> Brooke, Hammond, and Hirst, p. 234.

<sup>16</sup> M. DeForest, and E. Johnson, ‘The Density of Latinate Words in the Speeches of Jane Austen’s Characters’, *Literary and Linguistic Computing*, 16.4 (2001), pp. 389-401.

<sup>17</sup> *ibid.*, p. 389.

<sup>18</sup> *ibid.*

<sup>19</sup> A. DuBois, ‘Close Reading: An Introduction’ in *Close Reading: The Reader*, ed. by Frank Lentricchia and Andrew DuBois, (Durham: Duke University Press, 2003), pp. 1-40, (p. 2).

<sup>20</sup> *ibid.*, p. 5.

engage with the subtleties that shape the narrative. Close reading empowers scholars to discern thematic threads, character motivations, and cultural contexts that contribute to the broader literary field.<sup>21</sup> It encourages readers to engage with the text on multiple levels, teasing out layers of interpretation and contributing to a nuanced understanding of the work. It often involves exploring how literary devices function within the context of the narrative, considering the historical and cultural background, and even examining the author's personal experiences that might have influenced the creation of the text.<sup>22</sup> New critics such as Clench Brooks and I.A. Richards have however argued that close reading methods should focus exclusively on the text itself and that any knowledge about the author's biography should not be allowed to influence the reading and interpretation.<sup>23</sup> However, as DuBois notes, 'To understand the literary work, one must understand the cultural forms that help constitute it', which could also include biographical events.<sup>24</sup> Close reading invites readers to become detectives, uncovering the intricacies that give a text its unique character and depth. This research method functions as an essential tool in literary criticism.

While close reading focuses on a deep examination of individual texts, distant reading takes a broader and more quantitative approach. Distant reading, as illuminated by Verhaar, operates as both microscope and macroscope, enabling scholars to aggregate atomic details into multifaceted layers of analysis, uncovering patterns transcending individual textual units.<sup>25</sup> While close reading offers depth and intimacy with a single text, distant reading facilitates the exploration of a broader literary environment, offering a bird's-eye view of literary evolution and cultural shifts. Both techniques contribute uniquely to our understanding of literature by providing complementary perspectives, offering a balance between intensive exploration and extensive analysis. Close reading could potentially also be done computationally, but Eve predicts that 'when we do not understand the composition and collection of the underlying data sources on which we draw for comparative evidence, we can be led astray'.<sup>26</sup> Despite this risk, the two techniques could most certainly complement each other. As Verhaar notes, researchers confront the task of skilfully incorporating computation,

---

<sup>21</sup> *ibid.*, p. 14.

<sup>22</sup> *ibid.*, p. 38.

<sup>23</sup> I. A. Richards, 'The Analysis of a Poem' *Principles of Literary Criticism*, fourth edition, (London: Kegan Paul Trench, Trubner & Co LTD, 1930), pp. 114-33 (p. 114). SIDENOTE: Although Richards is saying this, he includes the nuance that the reader needs to be aware of the author's life to avoid misconceptions while reading but not use it as an explanation for analysis; Brooks, Cleanth, *The Well-Wrought Urn: Studies in the Structure of Poetry*, (London: Dobson, 1968), p. 165.

<sup>24</sup> *ibid.*

<sup>25</sup> Verhaar, p. 114.

<sup>26</sup> Eve, p. 112.

while also skilfully finding a wise equilibrium between thorough manual analysis and automated analysis, precision and comprehensiveness, and singular approaches versus multiple perspectives.<sup>27</sup> Distant reading complements traditional close reading and offers a complementary means to apprehend the broader currents shaping literary and cultural phenomena.

In the context of the research exploring the effects of textual analysis tools on literary scholarship in the digital age, the concepts of “close reading” and “distant reading” hold significant relevance. These approaches represent distinct methodologies in literary analysis. Each offering unique insights into the interpretation and understanding of texts. In this study’s case, close reading serves as a benchmark for traditional analysis, against which the outcome of digital tools can be evaluated. As distant reading employs computational techniques, it utilises digital tools for systematic analysis of large corpora of texts. In the context of Jane Austen’s works, distant reading, facilitated by textual analysis tools, could unearth previously unnoticed linguistic patterns or thematic variations across her novels which close reading could potentially miss. The distinction between close reading and distant reading is pivotal to this study as it forms the foundation for investigating the impact of technological advancements on literary analysis. By comparing the results and insights derived from these distinct methods, the research delves into the extent to which digital tools enhance or alter the understanding of Austen’s texts. This exploration ultimately contributes to a nuanced comprehension of how technology reshapes the act of reading and its implication for the broader field of literary studies.

### *1.5. Influences of Medium*

For this study, it is also important to understand the influence exerted by the textual medium on reading practices. The differences between reading a printed text or digital text and the ways in which they both offer an alternate reading experience are particularly relevant in this context. In determining how digital tools enhance the reading experience and literary analysis, these two mediums need to be explored more thoroughly. In Adriaan van der Weel’s exploration of changing reading practices in *Changing our Textual Minds*, he notes that the medium through which we read shapes and transforms our cognitive processes and interactions with texts. Printed text has long been the dominant medium for reading, and it has played a significant role in shaping traditional reading practices. The physicality of printed

---

<sup>27</sup> Verhaar, p. 151.

books, with their tactile sensations and linear narrative structures, has influenced how we engage with written content.<sup>28</sup> The act of turning the pages, the organisation of chapters, and the spatial layout of text all contribute to a structured reading experience.<sup>29</sup> This medium has encouraged deep, focused, and immersive reading practices that promote sustained attention and critical thinking.

However, the digital medium has brought about a paradigm shift in reading practices. Van der Weel's analysis points out several key influences of the digital medium on how we read. The first aspect to consider are multimodality and interactivity. Digital texts often incorporate multimodal elements such as images, videos, hyperlinks, and interactive features. These additions create a more dynamic reading experience, allowing readers to engage with content in multiple ways.<sup>30</sup> Another difference between the two media is spatiality and scrolling. Digital reading often involves scrolling through continuous streams of content, as opposed to the discrete page-turning of printed books. This scrolling behaviour can affect how we process information, potentially leading to skimming and shallower engagement.<sup>31</sup> It may also impact our ability to retain information and make connections between different parts of the text. Distractibility and fragmentation are also key influences of the digital medium, as this medium can encourage a fragmented reading experience due to the constant availability of distractions like social media, notifications, and ads. This can hinder deep, focused reading and contribute to shorter attention spans.<sup>32</sup> An additional important element van der Weel discusses is the possibility of annotation and collaboration in a digital medium. Digital tools often facilitate annotation, highlighting, and sharing of texts. This can promote collaborative reading practices and the exchange of ideas but may also introduce challenges related to information accuracy and context. As van der Weel puts it: the 'ease with which a digital text can be changed unfortunately goes hand in hand with the invisibility of the change. Thus, readers are hard-pressed to know which version of a text they are reading.'<sup>33</sup>

Van der Weel's insights highlight that the transition to digital reading practices is not simply a matter of adopting a new medium; it involves a transformation of cognitive processes and engagement with the text. Readers must navigate the tensions between the advantages of digital media (accessibility, interactivity, searchability) and the potential

---

<sup>28</sup> Weel, 'Salient Features of Digital Textuality', p. 183.

<sup>29</sup> *ibid.*, p. 170.

<sup>30</sup> Weel, 'Salient Features of Digital Textuality', p. 170.

<sup>31</sup> *ibid.*, p. 146.

<sup>32</sup> *ibid.*,

<sup>33</sup> *ibid.*, p. 180.

drawbacks (distractibility, shallower engagement, fragmentation). Essentially, the medium through which we significantly read shapes our reading practices and experiences. As digital media continues to evolve and become more integrated into our lives, it is important to consider how these changes impact our ability to deeply engage with content, critically analyse information, and maintain sustained attention.

These differences between printed and digital texts, and the impact these differences have on reading practices are important to consider when researching the impact of digital tools on literary scholarship. These differences are also of paramount relevance to this study investigating the implications of textual analysis tools for literary scholarship. In an era where digital platforms have become omnipresent, transforming the very nature of the engagement with texts, understanding these differences is crucial to comprehending the evolving territory of reading and its impact on literary analysis. Printed texts have a historical and palpable significance. Printed books often evoke a contemplative reading style, allowing readers to savour each word and engage in unhurried reflection. For traditional scholars accustomed to close reading, these physical and psychological dimensions play a pivotal role in their analytical process. On the other hand, the digital medium introduces a dynamic and interactive dimension to reading. However, this digital environment can also promote a more fragmented and distracted reading style due to the constant influx of notifications and the temptation to multitask.

In the context of Jane Austen's works, the choice of medium can impact the perception of her narratives and themes. The immersive, uninterrupted nature of print reading might allow readers to fully engage with Austen's intricate social commentaries and character developments. Conversely, the digital format might enable quick cross-referencing of characters and themes across her novels, potentially revealing new patterns that might have been overlooked through traditional close reading alone. By investigating the disparities between digital and printed reading experiences, this study delves into how technology-induced shifts in reading habits might influence the outcomes of literary analysis. Ultimately, understanding the multifaceted nature of reading experiences in various mediums enriches the study's exploration of how technology is reshaping the relationship between readers, texts, and the analytical process in the realm of literary scholarship.

#### *1.6. Structure of the Thesis*

The digital age has ushered in a profound transformation in the way we currently engage with literature, propelling us into a realm in which technology and literary scholarship intersect in

unprecedented ways. This juncture prompts an intriguing inquiry into the implications of these technological advancements for our comprehension of literary masterpieces. This thesis concentrates on the celebrated author, Jane Austen, whose nuanced narratives, and incisive social commentaries have captured the imagination of readers for generations. In this endeavour, I want to embark on a meticulous journey, employing computational textual analysis tools to add to the dimensions of Austen's literary universe, inspired by John Burrows.<sup>34</sup>

The study's foundation lies in the recognition that our literary interaction has undergone a profound evolution, with digital mediums altering the very fabric of how people read, analyse, and interpret texts. While engaging with this complex interaction of technology and literature, it's crucial to identify the various ways in which these shifts directly connect with and influence the detailed narratives found within Austen's novels. This thesis is comprised of two main chapters, one to establish a groundwork that explores existing theory on Jane Austen, which is primarily compiled through close reading. The other will be a case study conducted on Jane Austen through computational tools. The case study's core structure delineates five critical facets of Austen's literary craftsmanship, each scrutinised through the lens of computational analysis. Beginning with the exploration of diction and vocabulary, this study delves into the unique words that weave the tapestry of Austen's prose. By employing textual analysis tools, I will quantify her lexical distinctiveness, gauging the frequency of uncommon terms and analysing her linguistic choices. This not only offers a quantitative gauge of her lexical individuality but also underscores the potential evolutionary trajectory of her linguistic style across her canon.

The study then unravels the intricate threads of Austen's linguistic structure. Sentence lengths, complexities, and syntactical categories are systematically examined to ascertain the rhythmic cadence that underpins her storytelling. Furthermore, the study ventures into the realm of figures of speech, uncovering alliteration and internal rhyme through computational scrutiny. By mechanising the identification of these literary devices, it can unveil patterns that inform the understanding of Austen's narrative artistry. The realm of figures of thoughts – metaphors, imagery, and irony – beckon next. Though often the domain of close reading, this study will attempt to ascertain the feasibility of computational analysis in decoding these abstract elements. This promises not only an objective assessment of their prevalence but also unveils the potential nuances they contribute to Austen's storytelling.

---

<sup>34</sup> J.F. Burrows, *Computation into Criticism*.

In the final chapter of this computational exploration, this case study embraces the advance technique of topic modelling. With this, the essence of Austen's thematic preoccupations can be distilled, uncovering the latent patterns that traverse her narratives. This quantitative mapping of literary themes not only enriches our interpretation but also offers a glimpse into the thematic contours that might have eluded traditional analysis. By juxtaposing the digital tools' revelations with the cherished traditions of close reading, this study will decipher the intricate dance between technology and literary scholarship. This investigation not only enriches the current understanding of Austen's indelible mark on literature but also resonates as a microcosm of the broader evolution transpiring in the world of literary studies. In this convergence of the classical and the contemporary this study will unravel the timeless charm of Austen's narratives but also kindles a dialogue about the intricate interplay between technology, literature, and the art of interpretation.

## Chapter 2 – Reading Jane Austen in the Print Age

### *2.1. Introduction*

Within the bound pages of printed volumes, Jane Austen's literary legacy has flourished across centuries, transcending temporal confines to reach modern readers with timeless resonance. This chapter embarks on an exploration of Austen's novels within the traditional realm of the print age, tracing the contours of her genius as perceived through close reading and established literary criticism. In an effort to come to a conclusive answer to the main question as to how the digital medium and the utilisation of textual analysis tools affect the reading experience for literary scholars, and the relevance of these tools in literary criticism, it first needs to be established what scholars have discovered through traditional close reading methods.

The first subchapter delves into the perception of Jane Austen within the English literary canon, analysing her place among other authors and examining the criteria by which she is evaluated. It considers her reputation and the enduring significance of her works in literary discourse. The second subchapter investigates the genres that Austen employs in her writing and how she incorporated them. The third subchapter connects to this by exploring the thematic richness of Austen's novels, such as the societal, moral, and romantic themes that pervade her narratives. It traces the development of these themes across her works, revealing nuances and shifts in her exploration of human experiences. The fourth subchapter delves into Austen's characterisations and examines the techniques she employs to animate her characters. It analyses how she constructs and develops her characters, illuminating the complexities and nuances of their personalities. In the fifth subchapter a foundation will be created as to Austen's diction. It scrutinises her diction and vocabulary choices, exploring how they reflect social contexts and contribute to the portrayal of her characters and setting, through close reading methods. In the sixth subchapter a brief analysis of Austen's use of literary devices, such as subtle irony, imagery or intricate figures of speech will be given. It examines how these devices serve as vehicles for her wit, social commentary, and narrative craft. Concluding the chapter, the seventh subchapter investigates Jane Austen's enduring presence in popular imagination and her status within the 19<sup>th</sup> century literary canon. It explores the factors that have cemented her legacy and discusses the ongoing relevance of her narratives in contemporary literary discourse. All the analyses in the chapters are based upon the work of established scholars, experts in the field of Jane Austen. By basing it on decades worth of research, this chapter as a whole provides an important foundation to the

computational case study in the third chapter of this thesis. This chapter stands as a bridge connecting the past with the present, offering a panoramic glimpse into Austen's craftsmanship and the perennial allure of her narratives in the realm of the print age. By establishing all the above, a foundation is created to explore the relevance of digital tools, both for the reading experience and within literary criticism.

## *2.2. Views on Austen's position in the canon of English literature*

Within the literary discourse surrounding Jane Austen's novels, a complex tapestry of opinions and critiques has been woven. Her writing style, interwoven with societal themes, has garnered both praise and scepticism from scholars across the ages. Jan Fergus, in her exploration of the historical context in which Austen's works flourished, highlights the prevailing aristocratic disdain towards women writers' pursuit of print.<sup>35</sup> This disdain, rooted in the belief that such endeavours were vulgar, particularly afflicted elite women and some men. Fergus illuminates the societal hurdles that Austen and her female contemporaries confronted by saying; 'Only desperate financial need, preferably to support aged parents, a sick husband, or destitute children, could (according to literary men) excuse a woman's exposing herself in print to obtain money'.<sup>36</sup> This historical backdrop underscores the significance of Austen's own challenges as a female writer navigating an era defined by class and gender constraints.

Isobel Grundy, diverting the conversation towards Austen's own reading habits and influences, signals a shift from accepting Austen's claim of 'ignorance' at face value.<sup>37</sup> This evolving perspective encourages a more nuanced understanding of Austen's intellectual milieu. Grundy's stance resonates with recent scholarship that redefines Austen's intellectual influences, prompting questions about the intersection between Austen's creative individuality and her engagement with existing literary currents. Grundy's assertion of Austen's role as a bridge between literary epochs resonates across this discourse.<sup>38</sup> Her portrayal of Austen's narratives as a 'slender thread' carrying realism through the Romantic age is apt, but also provokes queries about the extent of Austen's deliberate influence on this transition.<sup>39</sup> Her work, says Clymer, is 'the slender thread which carried the strain of realism

---

<sup>35</sup> J. Fergus, 'The Professional Woman Writer' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 1-20 (p. 3).

<sup>36</sup> Fergus, p. 3.

<sup>37</sup> I. Grundy, 'Jane Austen and Literary Traditions' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 192-214 (p. 195).

<sup>38</sup> *ibid.*, p. 196.

<sup>39</sup> *ibid.*

safely through the Romantic age from the hands of Fielding and Richardson to those of Thackeray and Trollope'.<sup>40</sup> This analysis thus offers both an acknowledgment of Austen's continuity and a challenge to probe deeper into her conscious role in shaping literary trajectories. Thomas Keymer delves into the complex layers of revision present in Austen's works, challenging the idea that Austen's admission of 'belatedness' should be universally accepted as an absolute truth.<sup>41</sup> Keymer's insight into *Northanger Abbey* and *Sense and Sensibility* underscores Austen's dynamic and evolving creative approach. Regarding *Northanger Abbey* he notes that the 'novel also bears traces of events and texts that Austen could not have known when she wrote the original version, or even when she sold it to Crosby'.<sup>42</sup> Additionally, he argues that elements 'of parody and metafiction are less overt in *Sense and Sensibility*, and it has been argued that Austen at last broke into the world of publication by melting into the generic background rather than by asserting her difference'.<sup>43</sup> This perspective unveils a facet of Austen's creative process that computational tools could potentially illuminate further.

Jocelyn Harris delves more into Austen's stance on the dramatic embodiment in *Mansfield Park*, introducing a dimension of Austen's works that is often overshadowed.<sup>44</sup> Her interpretation of Austen's engagement with dramatic representation hints at an intentional narrative ambiguity. Regarding a scholarly dispute on Austen's representation of drama Harris notes that 'critics blame Evangelicalism for Austen's apparent disapproval of drama in *Mansfield Park*, and yet the family put on plays, she enjoyed seeing plays performed, and she was extremely familiar with drama, players and theatre practices'.<sup>45</sup> This perspective opens intriguing avenues for considering how technology could capture the nuanced interplay between Austen's narrative subtleties and her characters' intricacies. Additionally, Penny Gay offers an interpretation of Austen's harmonious blend of romantic ideals and trenchant social commentary, particularly in *Emma* and *Persuasion*.<sup>46</sup> Gay's analysis showcases Austen's dual exploration of the human heart and societal dynamics, saying that Austen 'elegantly brings

---

<sup>40</sup> *ibid.*

<sup>41</sup> T. Keymer, 'Northanger Abbey and Sense and Sensibility' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 21-38 (p. 22).

<sup>42</sup> *ibid.*, p. 23.

<sup>43</sup> *ibid.*, p. 31.

<sup>44</sup> J. Harris, 'Pride and Prejudice and Mansfield Park' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 39-54 (p. 39).

<sup>45</sup> *ibid.*

<sup>46</sup> P. Gay, 'Emma and Persuasion' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 55-71 (p. 69).

together the two planes of the novel: its consciousness of contemporary political and social reality, and its deployment of romantic folk tale and the poetic imagination'.<sup>47</sup> This perspective beckons computational scrutiny to discern how Austen crafts this intricate blend through her language and narrative structures. Margaret Anne Doody, in highlighting Austen's early fiction as a 'mock-pastoral world,' introduces a lesser-explored dimension of Austen's work.<sup>48</sup> Doody's analysis of self-gratification as a prevalent motif resonates with Austen's portrayal of human desires within societal confines. She notes that the 'desire for self-gratification prevails everywhere – including in the heroine modelled on Jane's best friend and devoted sister'.<sup>49</sup> This contemplation invites computational tools to unravel latent patterns of motifs woven throughout Austen's works.

Amidst this diverse spectrum, Austen's enduring popularity emerges as a unifying chord. Her novels, often portraying the nuances of everyday life, have captivated readers across ages. Austen's exploration of human relationships and her perceptive societal observations continue to strike resonant chords with modern readers, reaffirming her place as a literary leader. Her enduring popularity serves as a testament to her unparalleled ability to blend entertainment with profound insights, yielding a timeless appeal that spans eras. In sum, the interplay of these varied viewpoints not only refracts Austen's writing but also sparks lively exchanges within the scholarly arena. While these perspectives occasionally harmonise, they also inspire thought-provoking dissonances, inviting deeper explorations of Austen's legacy. This discourse encapsulates the intricate interplay of gender, class, intellectual engagement, and artistic creation within Austen's literary canon, breathing life into her works for successive generations of readers and scholars.

### 2.3. *Literary genres*

Within the realm of literature, Jane Austen's literary genius is most evident in her use of literary genres. Her comedic prowess shines through as she navigates between Gothic descriptions and commonplace actualities. In *Northanger Abbey*, she masterfully captures Catherine's novel-fuelled desires, blending humour and romance.<sup>50</sup>

---

<sup>47</sup> *ibid.*

<sup>48</sup> M. A. Doody, 'The Early Short Fiction' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp.72-86 (p. 79).

<sup>49</sup> *ibid.*

<sup>50</sup> *ibid.*

Equally interesting is Austen's mastery of genre awareness. Her ability to incorporate elements from various literary traditions give her works a timeless quality, one that seemingly effortlessly transcends their immediate time frame.<sup>51</sup> As Keymer highlights, Austen's astute self-consciousness about fiction genres and narrative techniques reveals an author keenly aware of the nuances and intricacies of literary creations.<sup>52</sup> Beside the Gothic genre, Austen's books include a wide range of genres, including Romantic, Comedic, and Neoclassical elements. Nevertheless, she is often categorised within the realms of proto-literary realism and satire. However, a critical perspective might argue that Austen's self-awareness and genre exploration extend beyond satire. They signify a deliberate navigation of societal constraints. While Keymer focuses on her sharp sense of genre weaknesses, one could also view Austen's engagement with various genres as a subtle act of defiance, a testament to her autonomy as a female author challenging the literary norms of her time.

*Sense and Sensibility*, on the other hand, provides a fertile ground for examining the interplay between the real and the romanticised. A theme which can also be found throughout most of her novels, such as *Pride and Prejudice* and *Emma*, but is most clearly exemplified in *Sense and Sensibility*. Keymer's analysis of Marianne's approach to daily life, seen as a lived-out novel of sensibility, paints a picture of a character deeply entrenched in her fantasies.<sup>53</sup> However, an alternative perspective could delve into how Austen intentionally exploits Marianne's perspective to both emphasise the Romantic genre's tendencies and comment on the implications of such romantic idealism in real-world scenarios.

These critical insights intertwine with Doody's perspective on Austen's revolutionary transformation. Doody's analysis showcases how Austen's evolution aligned her works with the changing literary era, a shift that ushered her into publish-ability.<sup>54</sup> This transition captures Austen's keen awareness of the evolving literary field and her ability to adapt her narratives to both honour tradition and embrace innovation. As we traverse Austen's thematic and generic landscape, Doody's observation of Austen's unique blend of eighteenth-century qualities with Romantic ideals emerges as a guiding beacon.<sup>55</sup> This synthesis underscores Austen's mastery in harmonising divergent literary currents, allowing her narratives to resonate with multiple generations. In Austen's narrative symphony, the myriad of themes

---

<sup>51</sup> *ibid.*

<sup>52</sup> Keymer, p. 27.

<sup>53</sup> Keymer, p. 32.

<sup>54</sup> Doody, p. 74.

<sup>55</sup> *ibid.*, p. 75.

and genres intersect, resulting in a tapestry that complicates categorisation. As we navigate this intricate landscape, the various perspectives offered by scholars weave together, enriching our understanding of Austen's multifaceted artistry. Each perspective invites us to unravel layers, inviting deeper contemplation of Austen's genius and her lasting impact on the literary canon.

#### 2.4. *Thematic concerns*

Venturing into Jane Austen's literary realm, one is immediately immersed in a tapestry of themes, each thread interwoven with her narrative genius. These themes, ranging from societal critique to the complexities of human relationships, showcase Austen's unique blend of wit and self-awareness that has captivated readers across centuries. Margaret Anne Doody's insights into Austen's stance on culture highlight a dichotomy present in many of her works.<sup>56</sup> Doody's perspective invites us to question whether Austen's scepticism about culture being anarchy reflects a rejection of societal norms or a keen commentary on their flaws. It encourages contemplation of how Austen's exploration of culture dissects the intricacies of societal structures rather than rejecting them outright. Furthermore, Doody's assertion that Austen's novels were not the appropriate home for social criticism or moral questioning challenges us to explore the nuanced ways in which Austen subtly wove societal commentaries into her narratives.<sup>57</sup>

In contrast, Beth Lau presents almost completely contrasting views, suggesting that Austen participated in the Romantic fascination with brother-sister incest which may challenge traditional interpretations.<sup>58</sup> While this perspective might initially seem discordant with Austen's genteel narrative, it invites an intriguing discussion about the deeper layers of her works.<sup>59</sup> Lau's insights about flawed marriages in Austen's novels invite contemplation of whether these portrayals serve as societal commentaries or explorations of human imperfections.<sup>60</sup> They also raise questions about the agency of Austen's heroines and heroes within societal constraints.<sup>61</sup> This perspective urges the reader and scholar to scrutinise Austen's seemingly idyllic relationships, probing the depths of desire and self-reflection that

---

<sup>56</sup> *ibid.*, p. 81.

<sup>57</sup> *ibid.*, p. 83.

<sup>58</sup> B. Lau., 'Placing Jane Austen in the Romantic Period: Self and Solitude in the Works of Austen and the Male Romantic Poets' *European Romantic Review*, 15.2, 2004, pp. 255-67 (p. 259).

<sup>59</sup> Lau, p. 259.

<sup>60</sup> *ibid.*, p. 261.

<sup>61</sup> *ibid.*

underpin them. It also prompts a reflection on the intricate interplay of love, individuality, and societal norms that defines Austen's characters' romantic journeys.

In a broader context, Austen's engagement with themes defies simplistic categorisation. Austen's exploration of themes consistently surpasses conventional boundaries, frequently defying simplistic characterisations. For instance, in *Pride and Prejudice*, she delves into the complexities of class divisions and social hierarchies, blurring the boundaries between the aristocracy and the middle class.<sup>62</sup> This thematic diversity inspires the question whether her themes served as vehicles for her complex critique of society? Through these perspectives, Austen's works emerge as nuanced narratives encompassing timeless themes, societal commentary, and individual agency.

Delving deeper into Jane Austen's literary mosaic, the interplay of themes and genre emerges as a riveting facet of her narrative finesse. As Jocelyn Harris aptly observes, Austen's works often bear distinct tones and characterisations that, when juxtaposed, weave a richer narrative fabric.<sup>63</sup> *Pride and Prejudice* and *Mansfield Park* stand as compelling examples. Austen employs theatricality in the former to unmask the deceptions of daily life, by using dramatically embarrassing scenes or social interactions. While in the latter, she embarks on a profound interrogation of the more literal phenomenon of plays and play-acting.<sup>64</sup> The intriguing interplay of perspectives, exemplified by Penny Gay, propels the analysis forward. Gay's insights delve into the nuanced portrayal of characters' inner lives. In *Emma*, Jane Fairfax assumes the enigmatic role of an alternative literary mode, contrasting with the comedic trajectory of the vivacious Emma.<sup>65</sup> Gay's observation invites the reader and scholar to contemplate how Austen utilises characters to illuminate various facets of her narrative spectrum. Gay's exploration extends to the comedic underpinnings that shape Austen's narrative landscape. The humour within the narrative, as Gay points out, often arises from the dynamics of the society depicted. Even Mr. Woodhouse's comedic consent to arrangements underscores Austen's ability to blend social reality with the optimistic world comedy.<sup>66</sup> This duality evokes questions about Austen's purposeful negotiation of societal norms and comedic resolutions.

---

<sup>62</sup> Doody, p. 83.

<sup>63</sup> Harris, p. 53.

<sup>64</sup> *ibid.*

<sup>65</sup> Gay, p. 61.

<sup>66</sup> *ibid.*, p. 63.

Margaret Anne Doody's insights on the fantasticating capacity of the mind further enrich the thematic exploration. The juxtaposition of desire and narrative progression, as she suggests, underscores the dynamic relationship between individual wants and societal norms.<sup>67</sup> Austen's characters often navigate this tension, reinforcing the interplay between individual desires and societal expectations that characterise her narratives. This interplay between individual consciousness and societal influences resonates with Gay's assertion that, in Austen's fictive world, societal reality holds more sway than the musings of sensitive souls.<sup>68</sup> This observation adds layers to our understanding of Austen's character dynamics, illuminating the ways in which the societal context shapes their inner lives. Moreover, Austen's evolving narrative landscape finds its top in *Persuasion*. Gay's perspective highlights its poetic essence and its alignment with Romantic views of nature's regenerative force.<sup>69</sup> This resonance underscores Austen's ability to navigate multiple genres, allowing her works to exist as reflections of the literary climate while maintaining their unique identity.

### 2.5. Characterisation

Jane Austen's writing in her novels goes beyond the simple selection of words and language. Her careful choice of diction and vocabulary serves as a powerful tool to illuminate the intricacies of character, perception, deception, and societal norms within her narratives. Austen's characters, who often serve as the primary lenses through which readers perceive the story's events, grapple with the dissonance between spoken words and underlying intentions. This strategic use of language not only encourages readers to engage actively with the narrative. As Jocelyn Harris notes, words like 'acting', 'adorned', 'affection', and 'appearance' also underscore the theme of masks and hidden motives.<sup>70</sup> Terms like 'disguise', 'misrepresentation', and 'pretension' further emphasise the need to unveil the truth behind appearances.<sup>71</sup>

In Austen's literary realm, characters emerge as vivid personas that carry the narrative's weight, embodying an excess of traits and qualities that transcend the pages. Her artistry in characterisations is evident in her strategic pitting of round characters against flat ones, a technique showcased in *Sense and Sensibility*. Through a unique form of free indirect discourse, Austen revitalises well-trodden plotlines, like those of self-effacing heroines

---

<sup>67</sup> Doody, p. 79.

<sup>68</sup> Gay, p. 65.

<sup>69</sup> Gay, p. 66.

<sup>70</sup> Harris, p. 43.

<sup>71</sup> *ibid.*

juxtaposed with their more ostentatious counterparts.<sup>72</sup> In *Sense and Sensibility*, the contrasting dynamics between the Dashwood sisters, Marianne, and Elinor, are striking. Marianne's overly expressive temperament serves as a foil for Elinor's composed demeanour, creating a poignant commentary on emotional restraint versus unrestrained sentimentality.<sup>73</sup> This nuanced portrayal suggests a more profound social critique inherent in Austen's narrative, one that extends beyond mere satire of the novels of sensibility.<sup>74</sup>

The theatrical element of Austen's characters also shines through in her ingenious character development. For instance, the case of Mr. Yates in *Mansfield Park* exemplifies Austen's exploration of dramatic artifice. Yates's transformation from impassioned to well-mannered mirrors, as Jocelyn Harris notes, Garrick's theatrical prowess, highlighting the intricacies of external presentation versus internal identity.<sup>75</sup> Similarly, Penny Gay offers the example that Frank Churchill's comment to Emma about Jane Fairfax's dancing takes on new layers of meaning with knowledge of Jane's hidden feelings, showing Austen's masterful use of subtle irony.<sup>76</sup> Furthermore, Bernard Paris's insights into character motivations further enrich the discussion. In *Mansfield Park*, Fanny's inadvertent captivation of Henry Crawford reveals Austen's psychological insight into both characters.<sup>77</sup> While Crawford's flirtatious nature encounters an unexpected challenge in Fanny, his attraction stems from her unique blend of virtues, revealing his own moral complexity.<sup>78</sup> This complexity is also evident in Crawford's admiration for Fanny, showcasing his moral sense but highlighting his struggle against temptation.<sup>79</sup>

Another example of Austen's characterisation is Emma Woodhouse's character journey in *Emma*, which exemplifies Austen's prowess in transformation. Paris's analysis of Emma as a character moving from pride to humility is astute. Emma's progression from self-aggrandisement to self-castigation, from delusion to self-knowledge, underscores the subtleties of Austen's character portrayal.<sup>80</sup> Similarly, Elinor Dashwood's resilience in *Sense and Sensibility* showcases Austen's deep understanding of moral dilemmas and social

---

<sup>72</sup> Keymer, p. 32.

<sup>73</sup> *ibid.*, p. 33.

<sup>74</sup> *ibid.*, p. 36.

<sup>75</sup> Harris, p. 47.

<sup>76</sup> Gay, p. 62.

<sup>77</sup> B. Paris, *Character and Conflict in Jane Austen's Novels: A Psychological Approach* (Detroit: Wayne State University Press, 1978), p. 26.

<sup>78</sup> *ibid.*

<sup>79</sup> *ibid.*, p. 31.

<sup>80</sup> *ibid.*, p. 70.

injustice. Paris's observation that Elinor's response to an imperfect world is to remain superior aligns with Austen's intention to present multifaceted characters grappling with complex realities.<sup>81</sup>

Interestingly, Austen's departure from traditional narrative expectations can be seen in *Lady Susan*. Janet Todd highlights the audacious treatment of the titular character's fate as a marked departure from conventional literary patterns.<sup>82</sup> Austen's willingness to break the mould, allowing Lady Susan to escape the tragic ends typically reserved for adulteresses and villains, raises questions about societal norms and literary experimentation within Austen's evolving voice.<sup>83</sup> Furthermore, the strategic use of narrative gaps for moral contemplation is a hallmark of Austen's skill. Janet Todd's observation that Austen entrusts moral conundrums to 'the World', implying a sophisticated audience beyond her characters, underscores the complexity of Austen's writing.<sup>84</sup> This technique not only adds depth to the narrative but also reflects Austen's awareness of her readers' role as active interpreters.<sup>85</sup>

Austen's evolution is also evident in her meticulous manuscript revisions. Todd's analysis of the insertions in *Emma* showcases a writer's struggle to convey a character's emotions and circumstances.<sup>86</sup> This vulnerability in Austen's drafting process contrasts with her confident prose, shedding light on the intricate interplay between authorial intention and textual outcome.<sup>87</sup> Lastly, Todd's observation that Austen's awareness of her own illness informs *Sanditon* reveals the confluence of personal experience and creative expression. Austen's exploration of invalidism mirrors her own physical struggles and provides a unique window into her mindset during its composition.<sup>88</sup> This insight into the intersection of Austen's life and literature adds depth to our understanding of her evolving authorial identity.

In short, Austen's characterisations are a masterful interplay of contrasting personalities, social commentary, and psychological depth. Through her characters, she navigates a wide spectrum of human emotions, exploring the intricate balance between appearances and realities, restraint and emotion, and pride and humility. Austen's characters

---

<sup>81</sup> *ibid.*, p. 185.

<sup>82</sup> J. Todd, 'Lady Susan, The Watsons, and Sanditon' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 87-96 (p. 88).

<sup>83</sup> *ibid.*

<sup>84</sup> *ibid.*, p. 89.

<sup>85</sup> *ibid.*

<sup>86</sup> *ibid.*, p. 91.

<sup>87</sup> *ibid.*

<sup>88</sup> *ibid.*, p. 92.

transcend the confines of her novels, becoming enduring archetypes that continue to captivate readers across generations.

### 2.6. Diction

Jane Austen's narrative craftsmanship extends to her choice of diction, revealing her ability to bridge the sensibilities of her era with timeless literary essence.<sup>89</sup> Keymer delves into Austen's narrative, showing her playful manipulation of Gothic tags within *Northanger Abbey*. Barbara MacMahon goes deeper into this phenomenon, arguing that the novel has obvious Gothic passages, when the diction and style of writing changes to use several 'body-part actors and goals'.<sup>90</sup> This aims to evoke a sense of excitement and curiosity in the reader. MacMahon provides examples for this, showing sentences such as:

“Catherine's heart beat quick”, “a cheek flushed by hope”, “an eye straining with curiosity”, “her fingers grasped the handle of a drawer and drew it forth”, [which suggests] both excitement and body-parts acting semi-independently of the (rational) mind. What the fingers find, eventually, is a roll of paper, “a striking exemplification of what Henry had foretold” (Austen, 135).<sup>91</sup>

This linguistic shift adds depth to the narrative, contributing to the thematic richness of the novel.

Moreover, the relationship between language and appearance is particularly evident in Mr. Knightley's action of displaying his mud-free shoes to Mr. Woodhouse in Austen's novel *Emma*. This seemingly mundane act carries symbolic weight, reflecting the broader societal emphasis on propriety and outward appearance.<sup>92</sup> The reference to mud becomes a metaphor for the imperfections and decay that lie beneath the veneer of social conventions. Margaret Anne Doody suggests that this portrayal signifies a sense of baulked relationships and sterility, revealing the complex interplay between language, propriety, and human relationships.<sup>93</sup> Austen's deliberate choice of words goes beyond mere linguistic precision; it contributes to the thematic depth of her novels. By using language to expose the tension between appearance and reality, Austen invites readers to join her characters in the quest to uncover hidden truths and navigate the complexities of social interactions. Through her

---

<sup>89</sup> *ibid.*

<sup>90</sup> B. MacMahon, 'Metarepresentation and Decoupling in *Northanger Abbey*: Part 2' *English Studies*, 90.6, 2009, pp. 673-94 (p. 676).

<sup>91</sup> *ibid.*

<sup>92</sup> Doody, p. 85.

<sup>93</sup> *ibid.*

mastery of diction and vocabulary, Austen crafts a nuanced linguistic world that mirrors the intricate world of her characters and the moral dilemmas they face. Aspects such as diction and language use, is where computational tools can add to the discussion, as they can perform more intricate analyses of this on written works.

As Jane Austen's literary journey unfolds, her diction and vocabulary evolve dynamically. Margaret Anne Doody aptly captures this evolution, noting how vitality infuses Austen's mature novels, where her 'religious feeling' resides.<sup>94</sup> This vitality in her language animates characters and situations with a palpable energy, making her works resonate deeply with readers. However, Doody also underscores the arid places where hypocrisy festers, hinting at the tension between the spirited and the spiritless in Austen's prose.<sup>95</sup> This transformation in Austen's literary style becomes more pronounced when we examine her early works, as Margaret Anne Doody suggests. These earlier fictions, characterised by an absence of the domestication that marks her mature novels, offer glimpses of an unbridled Austen.<sup>96</sup> The journey she undertakes to refine her style and voice is akin to a complex voyage, an intricate circuitous path she traverses to bring her unique regency dress back into fiction.<sup>97</sup>

Janet Todd's perspective on Austen's creative timeline sheds light on her evolution as a writer. Austen's progression is marked by diverse works, from the exuberance of her teenage years to the mature reflections of her final months. Todd's analysis underlines that Austen's career-spanning works provide a panoramic view of her creative processes.<sup>98</sup> However, this perspective also raises the question of whether Austen's consistent evolution was always linear, or if certain works represent deliberate returns to earlier stylistic choices.

Thus, Jane Austen's development as a writer is a tapestry woven with creative progression, experimentation, and the interplay between personal experience and fictional expression. Her diction and vocabulary shift and mature, reflecting dynamic interplay between societal expectations, her evolving voice, and the rich world of her characters.<sup>99</sup> This evolution underscores Austen's unique ability to resonate with readers across time, while also inviting us to explore the myriad facets of her narrative artistry.

---

<sup>94</sup> *ibid.*

<sup>95</sup> *ibid.*

<sup>96</sup> *ibid.*, p. 86.

<sup>97</sup> *ibid.*

<sup>98</sup> Todd, p. 87.

<sup>99</sup> Doody, p. 93.

### 2.7. *Literary devices*

In Jane Austen's literary realm, narratives are enriched by the deft employment of various literary devices. These tools, collectively referred to as literary devices, encompass an array of techniques that authors employ to convey ideas, evoke emotions, embellish their language, and engage readers. Such devices are a bridge between the writer's intentions and the reader's perception, creating layers of meaning within the text. As Franco Moretti suggests, examining these devices allows researchers to trace changes in their popularity across time, illuminating the diachronic and synchronic evolution of specific genres or techniques.<sup>100</sup> It is through this nuanced interplay between form and content that literary devices hold the power to imbue narratives with distinctive connotations and effects.<sup>101</sup> Within Jane Austen's novels, the strategic use of irony, imagery, figures of speech, and other literary devices serves as a multifaceted tool to convey subtle nuances, illuminate character traits, and add layers of meaning to the narrative. The effect of these devices is not unique to her writing, but how she employs them is. Austen's ingenious integration of these devices into her storytelling elevates her works to realms far beyond mere fiction. In *Mansfield Park*, for instance, Austen's manipulation of literary devices is evident through her ironic narrative style. The character Henry Crawford, mocked for his height by Mr. Rushworth, subtly echoes the stature of celebrated actors of Austen's era, such as David Garrick and Edmund Kean.<sup>102</sup> This astute parallel not only underscores Austen's keen understanding of the performing arts' complexities but also underscores the tension between surface appearances and hidden depths.<sup>103</sup>

Central to Jane Austen's narrative genius is her deft manipulation of irony, an artistry that often positions her as the ironical narrator. This irony is evident from the opening paragraphs of her works, as noted by Penny Gay. *Pride and Prejudice's* opening sentence, 'It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife', forms a case in points, and consequently became one of the most famous first lines in literary history.<sup>104</sup> Austen unabashedly takes on the role of an ironical narrator in these narratives, a stance that lays the foundation for her uniquely subversive storytelling.<sup>105</sup> This initial irony-laden encounter sets the stage for a narrative journey

---

<sup>100</sup> F. Moretti, *Distant Reading* (London: Verso, 2013), p. 47.

<sup>101</sup> Verhaar, p. 53.

<sup>102</sup> Harris, p. 45.

<sup>103</sup> *ibid.*

<sup>104</sup> J. Austen, *Pride and Prejudice*, (1813) Project Gutenberg, txt-file. Updated Jan 2021, p. 1.

<sup>105</sup> Gay, p. 55.

characterised by layers of interpretation and complex interplay between surface appearances and deeper realities. Austen's ironic nuances become apparent from her masterful narrative tailoring. Gay's observations of the 'ironical sting in the tail; captures the subtlety with which Austen employs ambiguity to guide reader interpretation.<sup>106</sup> This skilful crafting allows her to juxtapose the seemingly mundane with underlying ironies, challenging her audience to probe beneath the surface. A case point is Austen's narrative manipulation in *Emma*. Through Emma's lens, or more accurately, through the ironical narrator's gaze positioned just behind Emma's shoulder, Austen paints a vivid picture of irony in action. Emma's imaginative tendencies, emblematic of a novelist's perspective, frequently distort her view of reality. Austen capitalises on this, offering readers a window into Emma's creative and often misguided interpretations of her world.<sup>107</sup> This serves as a reflective pause, inviting contemplation on the art of novel-writing itself.

However, Austen's progression is not solely confined to societal critique; it mirrors her evolving narrative maturity. As Gay points out, Austen's art was progressing in a direction that allowed her to direct irony at various societal targets, signalling her increasingly robust satire.<sup>108</sup> In *Persuasion*, for instance, Austen turns her irony towards the snobbery and shallowness of characters like Sir Walter Elliot, whose predilection for the Baronetage reflects his self-absorption.<sup>109</sup> While Gay provides valuable insights into Austen's use of irony, a more nuanced understanding can be gleaned by considering the interplay between the ironical narrator and Austen's characters. The author's choice to employ irony through the lens of her characters not only adds depth to their portrayals but also accentuates the distance between their perceptions and reality. This deliberate misalignment serves as a critical commentary on the limitations of individual perspectives and underscores the need for readers to engage in a more nuanced interpretation.

Furthermore, while Gay acknowledges Austen's increasing satirical focus, it is worth noting that this progression doesn't merely extend to societal critique. It also mirrors Austen's evolving narrative maturity. As she delves deeper into the exploration of human foibles, her irony takes on a more multifaceted character, serving both as a lens to dissect social norms and as a mirror to reflect her characters' internal struggles. In this regard, Austen's irony transforms into a potent instrument of narrative revelation. However, Penny Gay's insights

---

<sup>106</sup> *ibid.*, p. 56.

<sup>107</sup> Gay, p. 56.

<sup>108</sup> *ibid.*, p. 64.

<sup>109</sup> *ibid.*

into Austen's narrative irony offer a valuable starting point in understanding the author's subversive storytelling. There are many more scholars who have delved into this aspect of Austen's writing, but for the purpose of this thesis it is not necessary nor vital to include a deeper exploration.

Austen's mastery of literary devices extends to her motif of books as more than mere objects. They become conduits for social interaction, debate, and imaginative experiences. Through characters' reactions to literature, Austen crafts a window into their psyches, revealing the profound influence of reading on their perceptions. This dynamic interaction between literature and life adds layers to her characters and to the narrative itself.<sup>110</sup> This interaction is woven into conversations, shaping characters' reactions to literature reveals her insightful grasp of how reading influences perceptions, from forming opinions on love and marriage to determining social conduct. This strategic interplay highlights Austen's ability to melt literature and life seamlessly, thus delving deeper into her characters' psychological landscapes. Austen's engagement with literary sources goes beyond direct imitation. Her subtlety lies in her avoidance of direct replication or heavy-handed allusion. She eschews self-congratulatory displays of erudition or adornment of her narratives with canonical epigraphs. Instead, her narrative strategy is more nuanced, often embedding the influence of literature in her characters' lives.<sup>111</sup>

For instance, in *Pride and Prejudice*, Burney's contribution of the novel's title resonates with the older generation, while Elizabeth and Darcy's pride and prejudice emerge from their personal experiences.<sup>112</sup> Austen's nuanced approach intertwines her characters' growth with her own narrative voice.<sup>113</sup> *Mansfield Park* serves as a canvas where Austen deftly navigates contemporary debates. Her familiarity with issues like the slave trade adds layers of context to her characters' dialogues.<sup>114</sup> Austen's ability to juxtapose societal concerns with personal experiences infuses her narrative with depth and relevance.<sup>115</sup> In *Emma* and *Northanger Abbey*, Austen's use of literary references reveals her characters' relationships with the world of ideas. Emma's reliance on bibliographies rather than direct reading mirrors her active nature.<sup>116</sup> Conversely, *Northanger Abbey* passionately defends

---

<sup>110</sup> Grundy, p. 193.

<sup>111</sup> *ibid.*, p. 194.

<sup>112</sup> *ibid.*, p. 209.

<sup>113</sup> *ibid.*

<sup>114</sup> *ibid.*

<sup>115</sup> *ibid.*

<sup>116</sup> *ibid.*, p. 210.

novels as a key literary form, employing irony to position them at the centre of the tradition. This balancing act highlights Austen's blend of critique, humour, and self-awareness.<sup>117</sup> Austen's manipulation of dramatic devices is equally noteworthy. In *Pride and Prejudice*, she deftly explores what Penny Gay terms the 'scopic seductiveness' of these devices and how they intricately complicate the very essence of perception.<sup>118</sup> This complex interplay of perspectives underscores her keen awareness of the mechanisms of human interaction, shaping the dynamics of her characters' relationships. Austen's unique engagement with the broader tradition of the novel is a testament to her literary depth.

In short, Austen's skilful use of literary devices and dramatic elements elevates her narratives beyond mere fiction. Her nuanced application of these tools adds layers of meaning, complexity, and resonance to her works, creating a literary legacy that continues to captivate readers across generations.

#### *2.8. Reception of Austen's work from the 19<sup>th</sup> century onwards*

In the 19<sup>th</sup> century literary canon, Jane Austen holds a distinguished position owing to her unique blend of intellect and satire. Her novels are not only literary masterpieces but also intricate reflections of the intellectual debates of her time. Austen's capacity to navigate these debates while maintaining a deliberate detachment from their direct influence underscores her subtle yet significant contribution to the literary landscape. As Isobel Grundy emphasises, Austen's aversion to being taken too seriously is matched by her keen awareness of the intellectual currents around her, leading to a nuanced and oblique engagement with the issues of her era.<sup>119</sup> This newfound perspective doesn't negate her own humorous self-deprecation but rather enriches our understanding of her literary significance. This distinctive approach is evident in her characters' interactions as well. For instance, the misinterpretation of clichés by characters like Mrs. Elton reflects their inadequacy as listeners, mirroring the way Thorpe and Mr. Parker fall short as readers.<sup>120</sup> Austen's incisive commentary on society extends to her portrayal of inadequate readers and listeners, who parallel the misreading of her work itself. This dual layer of meaning invites readers to transcend superficial understanding, encouraging critical engagement. The complexities that Austen embeds within her seemingly genteel narratives not only reflect her literary prowess but also resonate with a broader intellectual

---

<sup>117</sup> *ibid.*, p. 208.

<sup>118</sup> Harris, p. 41.

<sup>119</sup> Grundy, p. 197.

<sup>120</sup> *ibid.*, p. 198.

discourse of her time. Austen's adeptness at portraying both the art of discourse and the flaw in understanding further solidifies her status as a literary luminary whose insights and observations remain relevant, making her a revered figure in the 19<sup>th</sup> century canon.

Jane Austen's position within the 19<sup>th</sup> century literary canon is intricately woven with her engagement with her contemporaries. Her familiarity with the works of Samuel Richardson, for instance, is striking. Austen's extensive use of Richardson's novels is evident, as she draws from *Sir Charles Grandison* for *Pride and Prejudice*, later transposing its essence into a distinct key for *Mansfield Park*.<sup>121</sup> This level of intimacy with Richardson's works is demonstrated further by Austen's emulation of the heroine Harriet Byron's cap from *Sir Charles Grandison*. The presence of Richardson's *Clarissa* in the Steventon household further enriches her literary arsenal, as she reworked its elements in *Lady Susan* and *Sense and Sensibility*.<sup>122</sup> However, Austen's innovation goes beyond simple imitation. She introduces a unique narrative voice, saturated with a character's thoughts, reminiscent of an unspoken soliloquy, an element that she famously adds to Richardson's influence.<sup>123</sup> This blending of voices adds depth and complexity to her characters' inner worlds, painting a more intricate portrait of their feelings and motivations.<sup>124</sup>

Unlike her contemporaries, she abstains from overtly alluding to heathen mythology in her works. As Doody puts it, she 'is strikingly unlike her contemporaries, female as well as male, in not overtly alluding to any of the heathen mythology in her novels.'<sup>125</sup> Instead, she delves into the depths of the novel as a literary form, connecting with its profound traditions that stretch back through time. This approach allows her to escape the limitations of contemporary conversations and access more profound layers of storytelling. Austen's profound influences, while discreetly embedded, resonate strongly throughout her narratives.

While Austen engages with fellow novelists, she remains distinctive in her approach. Unlike other women novelists who engaged in explicit dialogues with Richardson, she stands a bit apart from the conventional paths of discipleship.<sup>126</sup> While her works engage with Richardson, Henry Fielding, and Frances Burney, they simultaneously reflect a consciousness of the limitations within each.<sup>127</sup> Her novels are not mere echoes but dynamic interactions that

---

<sup>121</sup> Harris, p. 40.

<sup>122</sup> *ibid.*

<sup>123</sup> *ibid.*

<sup>124</sup> *ibid.*

<sup>125</sup> Doody, p. 84.

<sup>126</sup> Grundy, p. 193.

<sup>127</sup> *ibid.*, p. 206.

critically assess and reshape existing literary traditions. In short, Jane Austen's place in the 19<sup>th</sup> century literary canon is defined not only by her use of contemporaneous influences but also by her unique approach to these influences. Her deep engagement with Richardson's novels, the infusion of unique narrative voices, and the manipulation of dramatic devices distinguish her as an author who both assimilates and innovates within her literary context, creating a legacy that goes beyond imitation to redefine the very essence of novelistic artistry.

In the vast literary canon, Jane Austen's enduring relevance is a testament to her unique ability to resonate with both scholars and amateur readers alike. Austen's literary allure isn't confined to any particular era but is rooted in her remarkable capacity to invite diverse interpretations, making her work a perennial subject of fascination and engagement. As Isobel Grundy suggests, Austen's value lies in the hands of her readers, who can shape and extract meaning from her narratives according to their varying perceptions and insights.<sup>128</sup> This intrinsic flexibility, coupled with Austen's incorporation of a wide range of ideas and her nuanced allusions, contributes to her lasting significance. Austen's novels embody intricate layers of thought and craft that invite scholarly scrutiny. Grundy stresses that Austen's work draws upon a multitude of texts and responses to them, forming a dynamic dialogue with her literary predecessors.<sup>129</sup> In this sense, Austen's interaction with her contemporaries transcends the boundaries of time, creating a tapestry of connections that offer scholars endless avenues for exploration. Analyses such as these are pursued in Deidre Lynch's terms, both by 'Janeites' and 'Austenian' readers. The former group consists of enthusiastic hobbyists and the latter group of dedicated scholars who view Austen as a point of intellectual inquiry and discourse.<sup>130</sup>

For scholars, Austen's narratives provide rich territory for analysis. Her works, while immensely entertaining, serve, as mentioned before, as vehicles for social critique, providing insights into the complexities of human nature and societal norms. In particular, her adept use of narrative voice as an unspoken soliloquy, as Harris points out, allows readers to delve into characters' thoughts and motives, opening the door for scholarly discussions on the nuances of her storytelling techniques.<sup>131</sup> However, Austen's influence extends well beyond the academic sphere. Juliette Wells emphasises the enduring appeal of Austen's novels for amateur readers, who engage with her works on their own terms, enjoying the narratives for

---

<sup>128</sup> *ibid.*, p. 211.

<sup>129</sup> *ibid.*

<sup>130</sup> D. Lynch, *Janeites: Austen's Disciples and Devotees* (Princeton: Princeton University Press, 2000), p. 12.

<sup>131</sup> Harris, p. 40.

their complexity as well as their entertainment value.<sup>132</sup> The ‘professional scholars/teachers/readers’ as Lynch characterises them, stand in contrast to the ‘Janeites’, highlighting the dual nature of Austen’s readership, each group contributing to her cultural relevance.<sup>133</sup>

Austen’s relevance in the 21<sup>st</sup> century can also be attributed to her seamless adaptation into popular culture. Scott Caddy highlights the phenomenon of Austen’s presence in pop culture through adaptations, expansions, and cross-references.<sup>134</sup> The most well-known adaptations are the BBC mini-series<sup>135</sup> of *Pride and Prejudice* from 1995 or the feature film inspired by the same book from 2005<sup>136</sup>. Both these examples are film adaptations that keep almost all the original material. However, there are also adaptations, such as the film *Bridget Jones’ Diary*, originally a novel, which are inspired by *Pride and Prejudice*, but are not literal adaptations of the novel.<sup>137</sup> Additionally, on ‘Archive of Our Own’ there are currently roughly 4500 works that are inspired by Austen’s novels.<sup>138</sup> Austen’s stories have become templates for various genres, reaching a broader audience through adaptations that range from traditional literature to multimedia platforms.<sup>139</sup> Her novels’ timeless themes of love, societal expectations, and individual growth continue to resonate with modern sensibilities, ensuring that Austen’s voice remains audible across changing times. Furthermore, the proliferation of adaptations speaks to the novels’ ability to seamlessly transition into different genres and forms while maintaining her essence. While scholars may initially dismiss these fictional adaptations as inventive departures from the original texts, they reveal much more about the imaginative ways readers wish to engage with Austen’s world.<sup>140</sup> The enduring popularity of Austen in various media forms attest to her universal themes and the universality of human emotions she masterfully portrays. In short, Austen’s enduring relevance in the 21<sup>st</sup> century rests on her ability to adapt to the evolving needs and preferences of readers. Her narratives remain open to interpretation, appealing to both scholarly exploration and personal enjoyment. Through a range of adaptations and cultural representations, Austen continues to

---

<sup>132</sup> J. Wells, *Everybody’s Jane: Austen in the Popular Imagination* (London: Bloomsbury Publishing, 2012), p. 64.

<sup>133</sup> Lynch, p. 113.

<sup>134</sup> S. Caddy, ‘Fan Media and Transmedia: Jane Austen in the Digital Age’ in *Fan Phenomena: Jane Austen*, ed. by Gabrielle Malcolm, (Bristol: Intellect Book, 2015), pp. 42-51 (p. 44).

<sup>135</sup> *Pride and Prejudice*, BBC Productions, (BBC, 1995).

<sup>136</sup> *Pride and Prejudice*, dir. by Joe Wright (UK, 2005).

<sup>137</sup> *Bridget Jones’ Diary*, dir. by Sharon Maguire (UK, 2001).

<sup>138</sup> Archive of Our Own, ‘Jane Austen’, <<https://archiveofourown.org/tags/AUSTEN%20Jane%20-%20Works/works>> (2 October 2023).

<sup>139</sup> Caddy, p. 49.

<sup>140</sup> Wells, p. 142.

captivate the popular imagination, serving as a bridge between scholarly discourse and the fervent enthusiasm of amateur readers.

## Chapter 3 – Reading Jane Austen in the Digital Age

### 3.1. Introduction

This case study presents a focused investigation of the impact of computational textual analysis tools on the interpretation of Jane Austen’s novels. Choosing Jane Austen as the focal point of this computational case study was a judicious decision owing to several compelling reasons. Firstly, Austen’s position as a literary icon and her enduring relevance makes her an ideal subject for explorations. Her literary works are characterised by keen social observations, intricate character dynamics, and refined wit, and they invite an in-depth analysis that computational tools can potentially enhance. Additionally, Austen’s works encompass a spectrum of narrative techniques, from the comedic irony of *Pride and Prejudice* to the introspective complexities of *Mansfield Park*. This diversity enables the study to evaluate the versatility of computational tools in capturing distinct literary nuances, answering the question *how do or could textual analysis tools enhance our understanding of Austen’s texts in the digital age?*

Austen’s novels have been meticulously examined through traditional literary analysis, offering a robust foundation for comparative assessment. This study can leverage existing scholarship as a touchstone while extending the exploration into new computational dimensions. Austen’s substantial body of work, and an additional 217 works of her contemporaries, provide a substantial corpus for computational analysis. This abundance of material ensure that the study can draw statistically significant conclusions, offering insights into overarching themes and subtle idiosyncrasies across her oeuvre. Her writing stretches both the end of the 18<sup>th</sup> and the beginning of the 19<sup>th</sup> century, embracing both Enlightenment rationality and Romantic sensibilities. By subjecting her texts to computational scrutiny, the study bridges these historical literary traditions with cutting-edge technology, creating an intersection of the traditional and the contemporary.

This meticulously crafted case study holds profound relevance within the evolving landscape of literary scholarship in the digital age. As technology continues to reshape the methods through which we engage with texts, this thesis seeks to illuminate the transformative potential of computational textual analysis tools within the realm of classic literature, exemplified by Jane Austen’s works. The term ‘classic literature; is often employed to denote works that have stood the test of time, possessing enduring relevance and resonance across generations. These texts are celebrated for their literary merit, cultural significance, and thematic universality, offering insights into the human condition that transcend temporal and societal boundaries. Classic literature encompasses a diverse array of genres, styles, and

themes, ranging from the timeless comedies of Shakespeare to the complex works of Tolstoj. What unites these disparate works is their capacity to provoke thought, evoke emotion, and encourage reflection on the fundamental aspects of human existence.<sup>141</sup> Through their rich narratives, complex characters, and incisive commentary, classic literary texts invite readers to contemplate profound questions about love, morality, identity, and society.<sup>142</sup>

However, while the canon of classic literature has traditionally been defined by a select group of authors and works, this classification has become under scrutiny in recent years. Critics have questioned the Eurocentric and patriarchal biases inherent in the traditional canon, advocating for a more inclusive and diverse understanding of what constitutes ‘classic’ literature. This evolving definition seeks to recognize the contributions of marginalised voices, women writers, and authors from non-Western cultures who have been historically overlooked and marginalised.<sup>143</sup>

By employing these tools to scrutinise Austen’s narratives, the study aspires to reveal layers of her artistry and shed fresh light on the intricacies that have captivated readers for centuries. The examination of unique words, linguistic structures, figures of speech, and thematic contours is poised to unearth formerly unnoticed patterns, lending an empirical dimension to our appreciation of her storytelling prowess. Moreover, this thesis goes beyond the singular scope of Austen’s works. It acts as a microcosm of the broader discourse on the convergence of technology and literary analysis. By contrasting the outcomes of computational analysis, of the type conducted by John Burrows in his monograph *Computation into Criticism*, for example, with traditional close reading approaches, the study aims to offer a nuanced understanding of how digital tools enhance, reshape, or challenge the interpretation of literary texts. In doing so, this case study endeavours to demonstrate the feasibility and value of computational methods in revitalising literary inquiry. Ultimately, this thesis hopes to illuminate a path toward a more holistic and dynamic future for the study of literature, where tradition harmoniously meets innovation.

### 3.2. *Computational Criticism on Austen*

In the realm of literary criticism, the advent of computational methods has ushered in a new era of analysis and interpretation. By utilizing digital tools, researchers have gained fresh insights into the works of authors such as Jane Austen. Before we delve into the computational case

---

<sup>141</sup> A. Rigney, ‘The Many Dimensions of Literature’, in *The Life of Texts: An Introduction to Literary Studies*, ed. by Kiene Brillenburg Wurth, and Anne Rigney. (Amsterdam: Amsterdam UP, 2019), p. 68.

<sup>142</sup> Rigney, p. 66.

<sup>143</sup> Rigney, p. 72.

study on Jane Austen's novels, we offer a brief introduction to the world of computational criticism as it pertains to Austen's novels. Together, Martin Paul Eve's *Close reading with Computers* and John Burrows' *Computation into Criticism*, form the main basis for understanding computational theory connected to literary studies and Jane Austen specifically. Paul Eve offers a foundational insight into the role of computers in literary analysis. He notes that computational methods do not replicate the way humans read; rather, they serve as tools to perform repetitive tasks, allowing human interpreters to focus on the results. This shift in labour distribution has significant implications for literary analysis.<sup>144</sup> One of the challenges that computational methods can address is the examination of different editions of text. These methods can identify subtle differences between editions, which can be valuable for close reading.<sup>145</sup> Additionally, computational, and digital visualisations enable us to understand the rearrangement of forms and help overcome the challenges of communicating textual modification in the context of contemporary copyright.<sup>146</sup> In the context of literary analysis, the term 'forms' encompasses various aspects of textual structure and composition. These include but are not limited to narrative structures, linguistic features, stylistic elements, and thematic patterns present within a text. Computational and digital visualisations provide invaluable tools for dissecting and comprehending these forms by offering visual representations of textual data.

Stylometry, a crucial facet of computational criticism, involves the use of computers to quantify the stylistic properties of texts. This practice has a rich history, from legal cases relying on stylometric evidence to authorship attribution.<sup>147</sup> John Burrow's delta, a well-regarded algorithm, plays a vital role in authorship attribution, even though the precise reasons for its success remain partially elusive.<sup>148</sup> Martin Paul Eve adds to this with the assumption that authors possess an unconscious aspect to their style, an aspect that is quantifiable and distinctive.<sup>149</sup> This perspective aligns with the notion that an author's subconsciousness can manifest itself via unique linguistic practices.<sup>150</sup> Burrows also emphasises the importance of statistical analysis and correlation in computational criticism.<sup>151</sup> These tools enable scholars to assess the significance of resemblances and relationships

---

<sup>144</sup> M. P. Eve, *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas* (Stanford: Stanford University Press, 2019), p. 2.

<sup>145</sup> *ibid.*, p. 43.

<sup>146</sup> *ibid.*, p. 51.

<sup>147</sup> *ibid.*, p. 62-3.

<sup>148</sup> *ibid.*, p. 66.

<sup>149</sup> *ibid.*, p. 69.

<sup>150</sup> Eve, p. 69.

<sup>151</sup> Burrows, p. 46.

among various elements in texts.<sup>152</sup> While computational methods offer powerful insights, they must be used in conjunction with traditional literary analysis to avoid biases and ensure a comprehensive understanding of the material.<sup>153</sup>

In the context of Jane Austen's works, computational criticism has shed light on several key aspects. One area of investigation is the evolution of her dialogue vocabulary. Burrows notes that while there are some changes in her use of words, common words tend to distinguish characters rather than mark a specific 'note' for each novel.<sup>154</sup> Characters will have a specific usage, and this is a pattern that can be found across the novels. For example, Burrows' analysis highlights that the word "quite" is used prominently by certain characters, especially women, in Jane Austen's novels. This word, with its historical emphasis on the absolute, reveals nuances in character dialogue. Darcy and Mr. Knightley, for example use "quite" differently, both main love interests.<sup>155</sup> Computational methods can enhance the understanding of Austen's novels by highlighting subtle details and patterns. Although they do not replace in-depth literary analysis, they serve as valuable tools to illuminate the literature and draw attention to aspects that might otherwise go unnoticed.<sup>156</sup>

Austen's portrayal of characters can also be scrutinised differently through computational analysis. Burrows points out that her more admirable characters tend to focus less on the past or the future, resulting in specific verb forms, such as 'very', 'of' and 'we', being less commonly used in their dialogue.<sup>157</sup> This reflects the complexity of her characterisations and the varied interpretations of characters like Marianne Dashwood.<sup>158</sup>

Computational criticism, as evidenced by the insights derived from the works of Jane Austen, provides a valuable lens through which to explore the world of literature. These methods offer new perspectives and help reveal the subtleties in an author's style and character development. By embracing the capabilities of computers in the study of literature, scholars can deepen their understanding of classic texts and marginalised texts and bring fresh awareness to their subtleties.

---

<sup>152</sup> *ibid.*, p. 57; p. 129.

<sup>153</sup> *ibid.*, p. 148.

<sup>154</sup> *ibid.*, p. 41-44.

<sup>155</sup> *ibid.*, p. 64.

<sup>156</sup> *ibid.*, p. 77-8.

<sup>157</sup> *ibid.*, p. 51.

<sup>158</sup> *ibid.*, p. 146.

### 3.3. Methodology

This case study mainly employs Book NLP or natural language processing (NLP) methods to analyse Jane Austen's works and that of her contemporaries computationally. Book NLP is a Python library created by David Bamman. It was crafted to navigate the intricacies and linguistic nuances of extensive textual works, particularly books and other lengthy documents. Originally, the project surfaced as a Java library in 2014, developed by David Bamman, Ted Underwood, and Noah Smith.<sup>159</sup> However, recognizing the burgeoning Python community within the digital humanities, Bamman transitioned Book NLP into a Python-based framework, enhancing accessibility and adoption among a broader spectrum of scholars and practitioners.

At its core, Book NLP is engineered to undertake a multifaceted analysis of textual content, encompassing critical NLP tasks such as part-of-speech tagging, dependency parsing, entity recognition, character name clustering, coreference resolution, quotation speaker identification, supersense tagging, and event tagging.<sup>160</sup> By leveraging sophisticated techniques rooted in machine learning and deep learning methodologies, Book NLP transcends the boundaries of conventional NLP libraries, offering a comprehensive suite of tools tailored to accommodate the unique challenges posed by large-scale literary works.

Unlike its predecessors which often grappled with the limitations of traditional NLP approaches when confronted with voluminous texts, Book NLP capitalises on cutting-edge technologies, namely the Python NLP library spaCy and the HuggingFace Python Transformer library.<sup>161</sup> This strategic amalgamation empowers Book NLP to harness the prowess of transformer models, renowned for their ability to decipher complex linguistic patterns and semantic structures with unparalleled accuracy and efficiency.

Using Book NLP as foundational method, stylometric theory is applied to discern the unique stylistic signatures of authors and genres with the corpus. Stylometric theory, as the name suggests, entails the quantitative measurement of stylistic properties within texts.<sup>162</sup> This approach has a rich and varied history, with application ranging from legal cases to authorship attribution. However, recent scholarship, such as that from the Stanford Literary Lab, challenges the notion that authorial signal is the sole determinant of stylometric

---

<sup>159</sup>Dr. W.J.B. Mattingly. 'Introduction to BookNLP' <[https://booknlp.pythonhumanities.com/01\\_intro.html](https://booknlp.pythonhumanities.com/01_intro.html)> (17 January 2024),

<sup>160</sup> Dr. W.J.B. Mattingly. 'Introduction to BookNLP' <[https://booknlp.pythonhumanities.com/01\\_intro.html](https://booknlp.pythonhumanities.com/01_intro.html)> (17 January 2024),

<sup>161</sup> *ibid.*

<sup>162</sup> Eve, pp. 62-63.

analysis.<sup>163</sup> Instead, there is recognition of additional signals, such as genre-specific features, that contribute to the stylistic fingerprint of a text.<sup>164</sup> By employing computational techniques, such as those implemented in Book NLP, stylometric analysis offers a novel perspective on literary texts.<sup>165</sup> These analyses could be described as a microscope or filtration process. Through this methodology, this study aims to identify not necessarily authorial signatures but more importantly genre signals and thematic motifs across the corpus of Jane Austen's works and those of her contemporaries.

While Austen's entire canon offers a rich tapestry of texts that can be analysed, this study will concentrate primarily on her six major novels: *Pride and Prejudice*, *Sense and Sensibility*, *Emma*, *Northanger Abbey*, *Mansfield Park*, and *Persuasion*. By delving into this selection, the study ensures a comprehensive yet manageable examination of Austen's diverse narrative styles, thematic explorations, and linguistic nuances. Additionally, a selection of contemporaneous authors is included for comparative analysis. These authors include Frances Burney, Charlotte Lennox, Hannah More, Charlotte Smith, Jane West, Mary Shelley, Emily Brönte, Charlotte Brönte and Anne Brönte, as they are her female colleagues writing in roughly the same time period. However, I have also included Samuel Richardson, Henry Fielding, William Makepeace Thackerary and Anthony Trollope as contemporaries, as it is often said that Austen was either inspired by them or formed an inspiration for them. Overall, these authors were chosen for a variety of reasons, including their thematic relevance, stylistic similarities, and historical context, thereby enriching the comparative framework of this study. All these authors' main works were included in the analysis, to create a full overview and comparison. The full list of titles is provided in Appendix A.

The data collection is facilitated through Jupyter Notebooks, each focussing on different aspects of literary analysis. These notebooks, developed in Python incorporate various computational techniques for text processing, character analysis, thematic identification, and lexical comparison. The data collected encompasses a wide range of textual features, including part-of-speech distributions, lexical richness, thematic motifs, and linguistic markers indicative of social critique and thematic engagement. The remainder of this chapter will present the results of the case study.

---

<sup>163</sup> *ibid.*, p. 70.

<sup>164</sup> *ibid.*

<sup>165</sup> *Eve*, p. 92.

### 3.4. Diction and Vocabulary

In this section I will delve into the results of the computational analysis focusing on Jane Austen's diction and vocabulary in her novels, specifically *Emma*, *Northanger Abbey*, *Persuasion*, *Mansfield Park*, *Pride and Prejudice*, and *Sense and Sensibility*. For this study, I analysed the concordances of specific words across Austen's novels. Each novel was examined individually, and words such as 'marriage', 'love', 'class', 'quite', 'society', and 'rank' were scrutinised for their occurrences. This analysis delved into the frequency of word usage, concordance, co-occurrence with other words, and proved insights into their contextual usage within Austen's works.

One of the key aspects of Austen's writing is her unique choice of words, which contribute significantly to the texture of her narratives. Unique words, those distinctively employed by an author, often serve as windows into their thematic and stylistic preferences. Let us explore some of the most notable findings. In Jane Austen's novels, the prominence to certain unique words sheds light on her narrative concerns. This was established through filtering her novels in a notebook dedicated to unique words, using Python, which flagged all unique words in Austen's novels.<sup>166</sup>

Working with Part of Speech tagging, it can be established that the most common nouns in the corpus include 'time', 'man', 'day', 'house', 'sister', 'family', 'room', 'friend' and 'mother'. These words clearly indicate a focus on domestic and interpersonal relations. The word 'marriage' occurs 210 times in the corpus. This relatively high frequency underscores the importance of the central theme of social and marital conventions, especially for women in the Regency era. Her persistent use of the word 'love', 480 times in total, reflects her exploration of romantic relationships, emphasising the importance of emotions and social norms. The word 'class' is notably scarce in her works. The word occurs only 14 times. This suggests that, although Austen does not address social distinctions elaborately, and that she chooses to focus on personal relationships instead. 'Society' and 'rank' are more common terms, with 149 and 41 occurrences respectively, reflecting her keen observation of societal structures and manners.

The novel *Emma* has the highest frequency of the word 'marriage' can be observed, namely 33 times. Austen discusses people's marriages, making this a central theme to the narrative. The connection between 'marriage' and other words like 'love' (110 times) and 'class' (5 times) reflect the interplay of these themes. In comparison, in *Northanger Abbey* the

---

<sup>166</sup>GitHub 'Austen in the Digital Age'. < <https://github.com/MMvVeenen/austen-in-the-digital-age> > (14 June 2024).

word 'marriage' occurs 11 times, primarily in the context of general expectations surrounding marriage. Interestingly, 'class' (1 time) and 'rank' (1 time) are scarcely mentioned, emphasising the novel's focus on young love. *Persuasion* on the other hand, even though it is the shortest novel in Austen's canon, has 27 mentions of 'marriage' which highlights its centrality in this novel. The connection to specific characters like 'Mr. Elliot' and 'Anne' demonstrates the intricate relationships Austen builds within the narrative. The word 'class' appears 4 times, further emphasizing the role of societal rank. In *Mansfield Park* the word 'marriage' appears 30 times, with specific connection to names like 'Fanny' and 'Thomas'. This illustrates Austen's emphasis on marital relationships and the influence of class, as evidenced by the occurrence of 'class' (3 times) and 'society' (30 times). In *Pride and Prejudice* the word 'marriage' features most prominently, 66 times, and is intertwined mostly with characters such as 'Lydia' and 'Elizabeth'. 'Love' occurs frequently (90 times) and is linked to numerous other terms, reflecting the centrality of romantic relationships in the story. The mentions of 'class' (1 time), and 'rank' (12 times) highlight the societal aspects central to the novels plot. Lastly, *Sense and Sensibility* mentions 'marriage' 43 times, primarily concerning people's marital affairs, with connections to characters like 'Elinor' and 'Lucy'. 'Love' (74 times) is a pervasive theme, closely connected to characters such as 'Marianne' and 'Elinor'. Although the term 'class' is not explicitly used, the word 'society' (24 times) indicates discussions about societal rank and social interactions.

In comparison to her contemporaries, Austen uses Words like 'sister' also achieve a high uniqueness value, indicating the significance of siblings in her narratives. While the words 'lord' and 'thou' are used by Austen's contemporaries, these words interestingly are never used by Austen. That 'lord' is not a unique word for Austen but rather for her contemporaries is somewhat surprising as this word has heavy connotations to rank and social standing. Instead, Austen's focus is more on the female standing rather than male.

Another interesting aspect I tested against is the use of the word 'quite'. John Burrows argued that this word is specifically used for and by women in Austen's novels.<sup>167</sup> Although this is mostly true according to my own tests, I found that for *Persuasion* this is not quite the case. In this novel, it is used to describe women, but in speech it is equally used by both female and male characters. Additionally, although 'quite' is used more by and for women in the other novels, it is not exclusively so as for instance in *Mansfield Park*, *Pride and*

---

<sup>167</sup> Burrows, p. 64.

*Prejudice*, and *Sense and Sensibility*. This is interesting because John Burrows' methods are the base which is used by many scholars to measure against or start from.

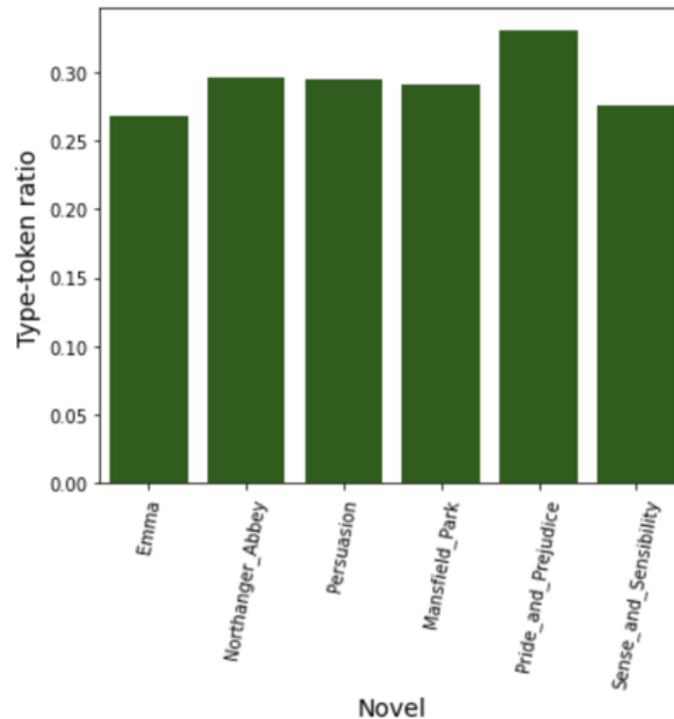


Figure 1. Type-token ratios of all novels

Another tool to quantify Austen's diction and vocabulary is to measure the Type-Token ratio. The type-token ratio measures the richness of an author's lexicon. In Austen's works, this ratio is relatively stable, as can be seen in figure 1. However, it is interesting to note that *Pride and Prejudice* has a slightly higher type-token ratio compared to her other novels. This suggests that *Pride and Prejudice* exhibits a more varied vocabulary, which aligns with its complex characters and social intricacies and could also hint at the reason why this is her most successful and popular novel. Compared to her contemporaries Austen's type-token ratio is relatively low. This implies that Austen's novels have a more repetitive vocabulary than the novels written by her contemporaries. Works such as *Camilla; or A Picture of Youth* by Fanny Burney and *Burlesques* by William Makepeace Thackeray exhibit a significantly higher lexical richness. As can be seen in figure 2, works like *The Wise Mamma Goose* and *How Freckle Frog Made Herself Pretty* by Charlotte Brontë display a lower type-token ratio.

### 3.5. Linguistic Structure

Analysing the linguistic structure of Jane Austen's works and comparing her to her contemporaries provides fascinating insights into her writing style and how it evolved. In this section, I will dive into the results and their implications.

The examination of the average sentence lengths (as measured in the number of words) across Austen's novels reveals a stable pattern, as is evident in figure 3. However, *Sense and Sensibility* stands out with longer sentences compared to her other works. This might reflect a more complex narrative structure or thematic depth in this specific novel. As this was one of her initial works, she might have felt inclined to provide a deeper exploration of themes and characters. Austen's sentence length is on the longer end of the spectrum among her contemporaries, as is evident in figure 4. While most authors of her time fell within her range, exceptions like Fanny Burney's *Cecilia* and Henry Fielding's *The Journal of a Voyage to Lisbon* featured longer sentences. In contrast, shorter sentences were notable in works like William Makepeace Thackeray's *The Wolves and the Lamb*.

The use of modal verbs, which were analysed using Part of Speech tagging, are quite stable, although *Pride and Prejudice* exhibits a higher usage of them, as can be seen in figure 5. This may indicate a more nuanced exploration of characters' inner thoughts, desires, and motivations in this novel. As one of her first novels, it is not strange that *Pride and Prejudice* diverges from the rest as this could have been Austen trying to establish her writing style. Austen's use of modal verbs is within the middle range compared to her contemporaries. Notably, Anthony Trollope with his novels *He Knew He was Right*, *The Last Chronicle of Barset* and *The Way We Live Now* has the highest use of modal verbs compared to all contemporaries. In contrast, Charlotte Brontë with *The Wise Mamma Goose* and *How Freckle Frog Made Herself Pretty* and William Makepeace Thackeray with *John Leech's Pictures of Life and Character* and *The Mahogany Tree* exhibit the lowest use of modal verbs.

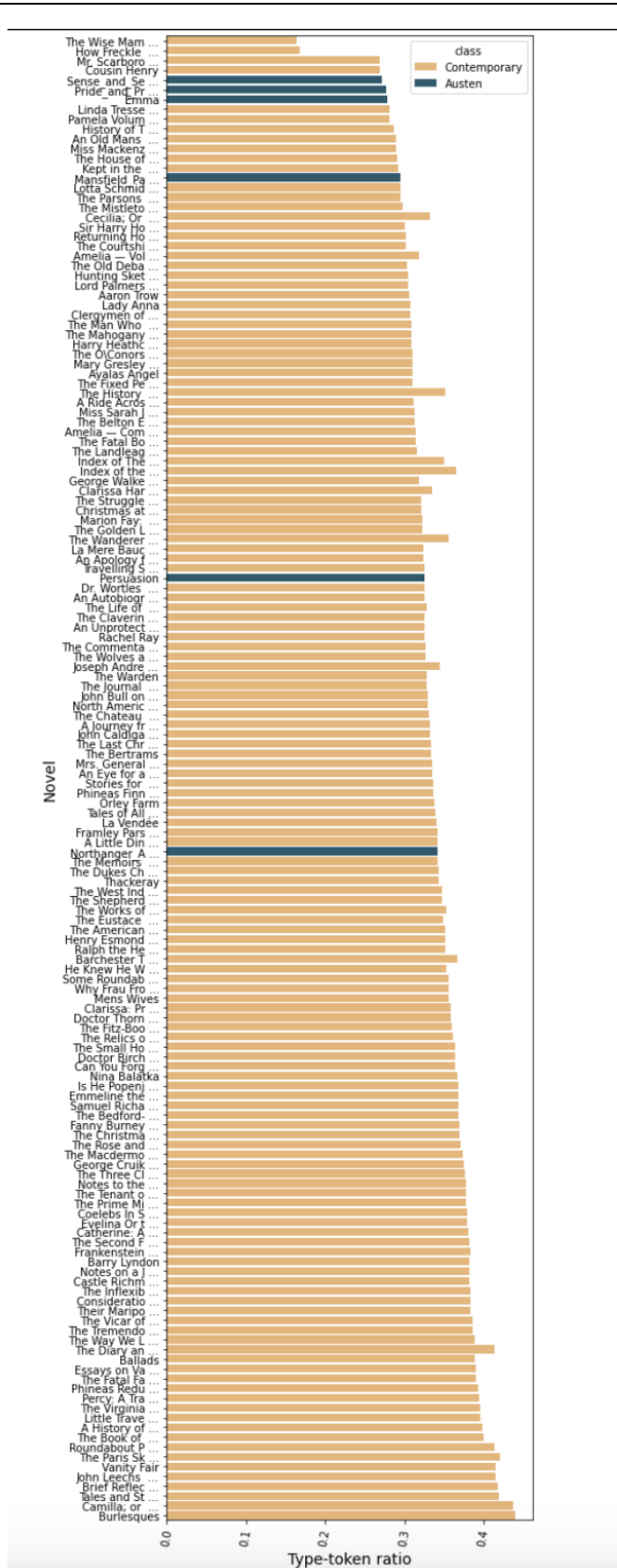
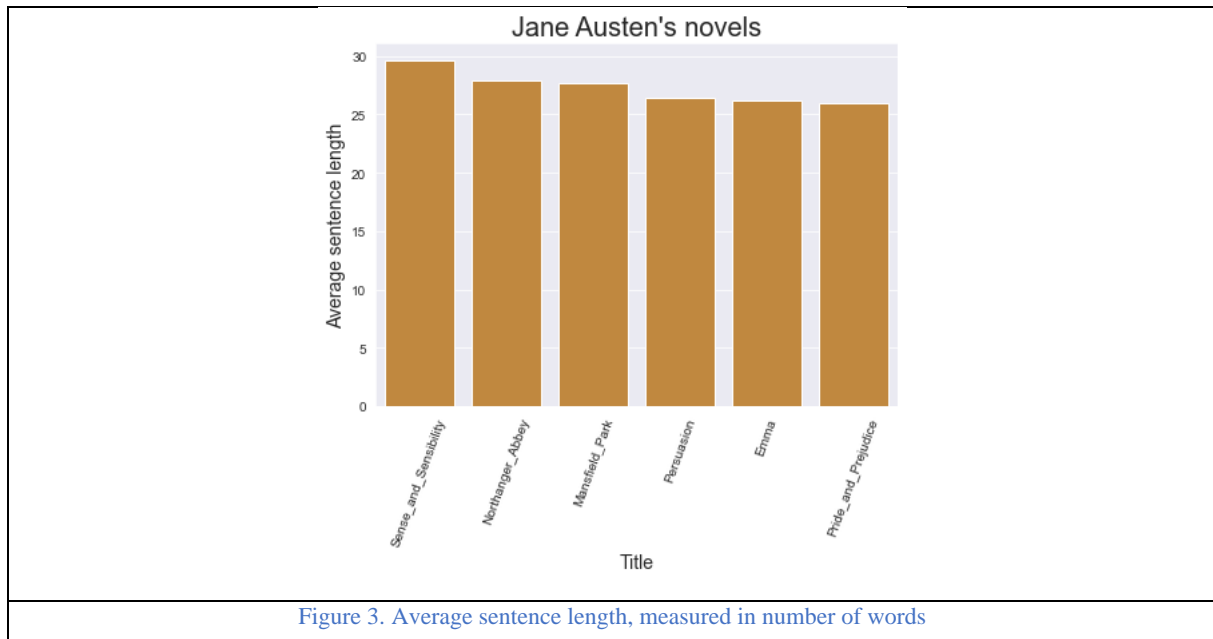


Figure 1. Type-Token Ratio Austen and Contemporaries



Additionally, analysing adjectives and adverbs, it is noticeable that all of Austen's novels exhibit a similar pattern, except for *Mansfield Park*, which stands out with a unique structure, as can be seen in figure 6. The variation in *Mansfield Park* suggests a distinct use of descriptive language, possibly contributing to the novel's specific tone or atmosphere. When looking at all her novels individually, it is interesting to see that their overall patterns are quite stable. Austen and her contemporaries share a common pattern in the use of adjectives and adverbs, though some novels deviate from this trend. While we can't specifically identify these deviations, they likely contribute to unique stylistic choices in those works.

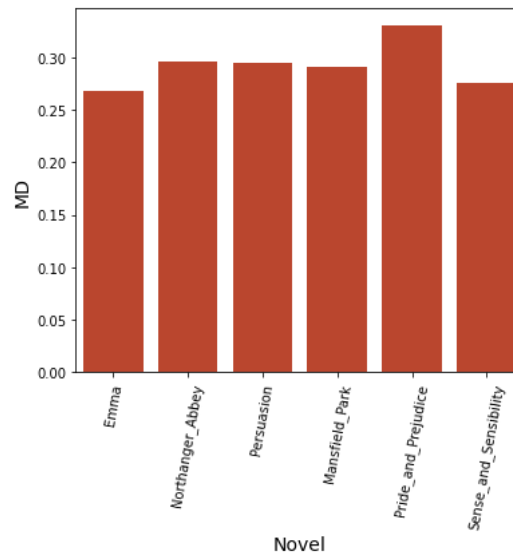
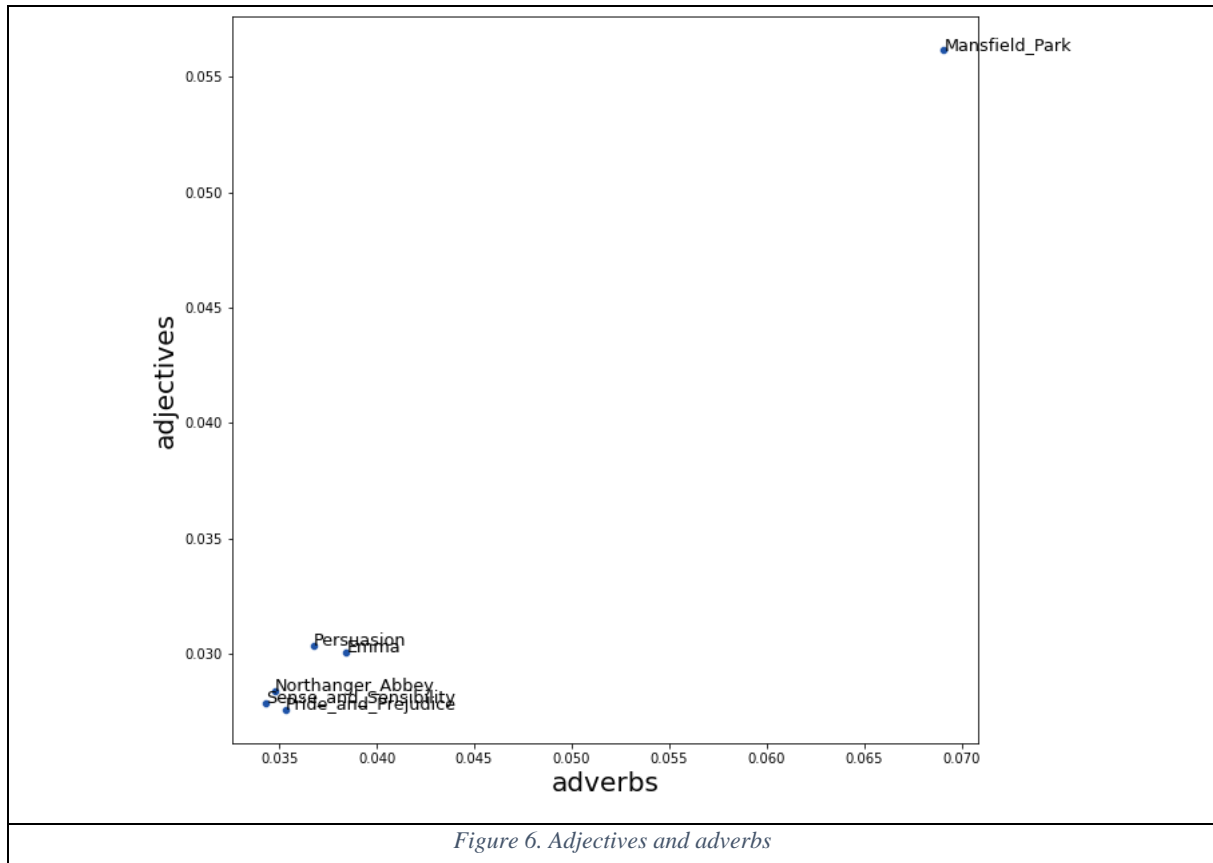


Figure 5. Use of modal verbs and adjectives

Austen's use of syntactical categories is extensive and diverse, showcasing her ability to create rich and layered narratives. The frequent appearance of categories related to characters, actions, and dialogue highlights the centrality of character development and interpersonal interactions in her novels. In interpreting these results, we can discern Austen's distinctive style. Her longer sentences and rich use of modal verbs imply a depth of character exploration and complex narratives. The consistent use of adjectives and adverbs across her novels suggests a coherent stylistic approach, with exceptions in *Mansfield Park* indicating unique linguistic choices tailored to that particular story.



Her distinctive style, marked by quantifiable elements such as sentence length, modal verbs, and syntactical richness, contributes to the enduring appeal to her novels. This is not only relevant for the literary scholar studying Austen, but also for the ‘professional’ reader, who after conducting or reading a computational study, could never read her works the same and will read her works with a new and profound understanding of the technical genius of her writing.

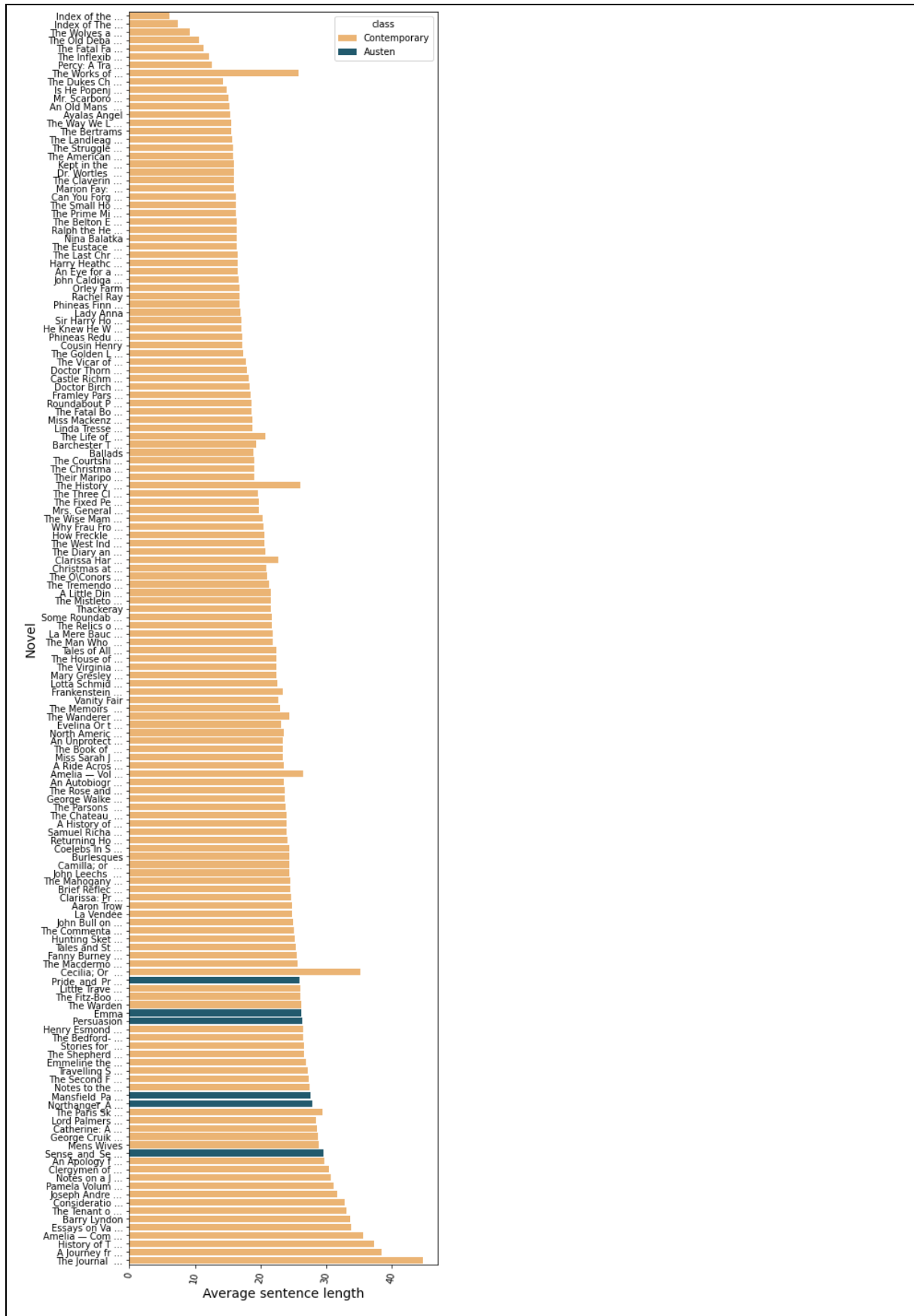


Figure 4. Average sentence length in 19th century novels.

### 3.6. Figures of Speech

In literary studies, exploring an author's use of poetic elements like alliteration and internal rhyme can provide unique insights into their style and thematic choices. Jane Austen, renowned for her wit and precision with language, employs these elements to enhance her narratives. In this case study, I tried to analyse these elements computationally, and will delve into these results and what they reveal about Jane Austen as a writer and computational literary studies. For alliteration, the computer looked for the same sound at the beginning of the following or closely connected word.<sup>168</sup> For internal rhyme, the computer looked for 'a poetic device by which two or more words rhyme within the same line of verse'.<sup>169</sup>

Alliteration and internal rhyme, while more commonly associated with poetry, appear to play a role in Jane Austen's novels, enhancing her narrative style and contributing to the musicality of her prose. I will discuss per novel the findings of this computational study, and while alliteration was relatively easy to find, internal rhyme proved to be difficult to identify. This has two possible explanations, the first being that Jane Austen did not employ internal rhyme regularly or deliberately in her novels, and the second being that this is not something that can easily be found in novels, as it can be in poetry.

Looking at *Mansfield Park*, many results surfaced which the computer picked up as examples for alliteration and internal rhyme. For instance, in the sentence 'She has Rushworth feelings, and Crawford feelings, and in the vicinity of Sotherton, the former had a considerable effect', we see the alliterative use of the 'f' sound in 'feelings' and 'former'.<sup>170</sup> This technique adds a lyrical quality to the prose and draws attention to the emotional nuances of the characters, making the narrative more engaging. Additionally, an internal rhyme example was found in the fragments "As far as walking, talking, and contriving reached, she was thoroughly benevolent" and "an answer, which comprehended each sister in its bitterness".<sup>171</sup> The correspondence in the final phoneme sequences in 'sister' and 'answer' in the latter example endow the sentence as a whole with a harmonious ring.

Moving on to *Sense and Sensibility*, this novel exhibits a subtle but purposeful use of alliteration. For instance, in the sentences "They gave themselves up wholly to their sorrow, seeking increase of wretchedness", and "still she could struggle, she could exert

---

<sup>168</sup> Oxford Dictionary of Literary Terms 4<sup>th</sup> ed. 'Alliteration' < <https://www-oxfordreference-com.ezproxy.leidenuniv.nl/display/10.1093/acref/9780198715443.001.0001/acref-9780198715443-e-31?rsk=fhORwt&result=1> > (1 June 2024).

<sup>169</sup> Oxford Dictionary of Literary Terms 4<sup>th</sup> ed. 'Internal Rhyme' < <https://www-oxfordreference-com.ezproxy.leidenuniv.nl/display/10.1093/acref/9780198715443.001.0001/acref-9780198715443-e-603?rsk=ETBkA0&result=1> > (1 June 2024).

<sup>170</sup> Austen, *Mansfield Park*, (1814) Project Gutenberg, txt-file. Updated March 2018, p. 142.

<sup>171</sup> *ibid.*, p. 19.

herself”, the repetition of the ‘s’ sound contributes to a smooth and harmonious flow of words.<sup>172</sup> This stylistic choice enhances the emotional depth of the narrative, showcasing Austen’s ability to use alliteration to convey evolving sentiments of the characters and create an engaging reading experience. As for internal rhyme, examples that were found include ‘[t]he late owner of this estate was a single man” and “you have been wholly engrossed on the most affectionate principle”.<sup>173</sup>

In *Northanger Abbey*, an example of alliteration can be seen in the sentences “she was now seen by many young men who had not been near her before” and ‘my dearest Catherine, what have you been doing with yourself all this morning?’.<sup>174</sup> This latter example, where the alliteration is on the words ‘dearest’ and ‘doing’, draws attention to the warmth and affection between Catherine and her friend. It underlines the close relationship between the two characters and the comradeship. The sentence ‘She listened reluctantly, and her replies were short’ contains an intriguing use of alliteration with the repetition of the ‘l’ sound in ‘listened’, ‘reluctantly’ and ‘replies’.<sup>175</sup> This repetition serves to emphasise the character’s hesitancy and succinctness, contributing to the overall rhythm of the prose.

As for *Persuasion*, the example ‘When he came away, she was going to walk out with Captain Benwick, which, he hoped, would do her good’, demonstrates Austen’s artful use of alliteration, where words like ‘walk’, ‘with’ and ‘would’ weave a subtle, rhythmic pattern.<sup>176</sup> In using alliteration as such, it adds a layer of auditory beauty to Austen’s prose, allowing readers to savour the language and emotions within the story. The example that the computer found for internal rhyme is: ‘Then might she again take up the book of books with as much enjoyment as in her early youth, but now she liked it not’.<sup>177</sup> The internal rhyme of ‘might’ and ‘liked’ marks a shift in Catherine’s relationship with reading. It serves as a subtle device to showcase her character development as she matures, and her tastes evolve.

In *Emma*, there is a well-crafted instance of internal rhyme to be found in the sentence ‘On the contrary, her plans and proceedings were more and more justified and endeared to her by the general appearances of the next few days’.<sup>178</sup> The internal rhyme is evident in the repeated ‘d’ sound in ‘proceedings’, ‘justified’, and ‘endear’. It adds a pleasant rhythm to the prose, subtly enhancing the reader’s engagement with the narrative. Examples

---

<sup>172</sup> Austen, *Sense and Sensibility*. (1811) Project Gutenberg, txt-file. Updated Dec 2020., p. 315.

<sup>173</sup> *ibid.*, p. 14.

<sup>174</sup> Austen, *Northanger Abbey*, (1817) Project Gutenberg, txt-file. Updated Jan 2021. p. 56.

<sup>175</sup> Austen, *Northanger Abbey*, p. 141.

<sup>176</sup> Austen, *Persuasion*, (1817) Project Gutenberg, txt-file. Updated Feb 2015. p. 203.

<sup>177</sup> *ibid.*, p. 12.

<sup>178</sup> Austen, *Emma*, (1815) Project Gutenberg, txt-file. Updated March 2018. p. 120.

of alliteration in Emma can be found in the sentences “It is impossible that Emma should not miss such a companion,” said Mr. Knightley.”, “Her father fondly replied”, “But there may be pretty good guessing” and “You know nothing of drawing”. In the sentence ““Of the lady, individually, Emma thought very little””,<sup>179</sup> the alliteration connected to ‘lady’ and ‘little’ underscores Emma’s attitude towards the lady in question, portraying her as a character who may not be as central to the narrative as others.

Finally, *Pride and Prejudice* has interesting examples of internal rhyme, such as the sentences “there is not another woman in the room whom it would not be a punishment to me to stand up with” and “there is one of her sisters sitting down just behind you, who is very pretty, and I dare say very agreeable”.<sup>180</sup> In the second example, the internal rhyme is woven very subtly into the text. As for alliteration, in the sentence ‘you may well warn me against such evil’, we can observe alliteration in the repeated ‘w’ and ‘m’ sounds in ‘may’ ‘well’ ‘warn’ ‘me’.<sup>181</sup> This alliteration adds a musical quality to the prose, underscoring the weight of the warning issued in this dialogue.

Alliteration and internal rhyme in Jane Austen’s novels serve to create a subtle musicality in her prose. These poetic elements enhance her storytelling and characterisation, contributing to the overall enjoyment of her works. They create a rhythm that draws readers into the narrative, making the experience more engaging. Additionally, these poetic elements can be seen as tools for emphasising emotions, character traits, or key moments in stories. Analysing alliteration and internal rhyme computationally can help scholars identify patterns and trends across an author’s body of work. This approach can uncover connections between sound devices and narrative themes or character development that might be less evident through traditional close reading.

### 3.7. *Semantic Categories*

Austen’s novels are renowned for their rich use of language and imagery, which contribute to the reader’s deep engagement with her characters and their emotions. While the computational analysis of metaphors and irony remains a challenging endeavour, we can gain valuable insights into her works by exploring word groupings related to emotion, plants, animals, cognition, and phenomena in her novels. These words were found using the

---

<sup>179</sup> *ibid.*, p. 310.

<sup>180</sup> Austen, *Pride and Prejudice*, p. 10.

<sup>181</sup> *ibid.*, p. 542.

Supersense Entities methods, available in the BookNLP package. Graphs per category were produced, providing insights into the distribution of these entities within Austen's works.

Emotion plays a central role in Austen's novels, and my analysis of words related to emotion in her works reveals a wide range of sentiments experienced by her characters. From feelings of joy and cheer to moments of panic and surprise, Austen's characters navigate a complex array of human emotions. This data, underscores the depth and complexity of Austen's characters and their emotional lives, providing literary scholars and readers with a better understanding of the nuances in her storytelling. For instance, in *Pride and Prejudice*, the character Elizabeth Bennet experiences a range of emotions, such as prejudice, pride, love, and eventually, self-awareness. This novel is an excellent example of how emotions and personal growth are central themes in all Austen's works. Her characters often undergo emotional journeys, reflecting the complexities of human feelings in society.

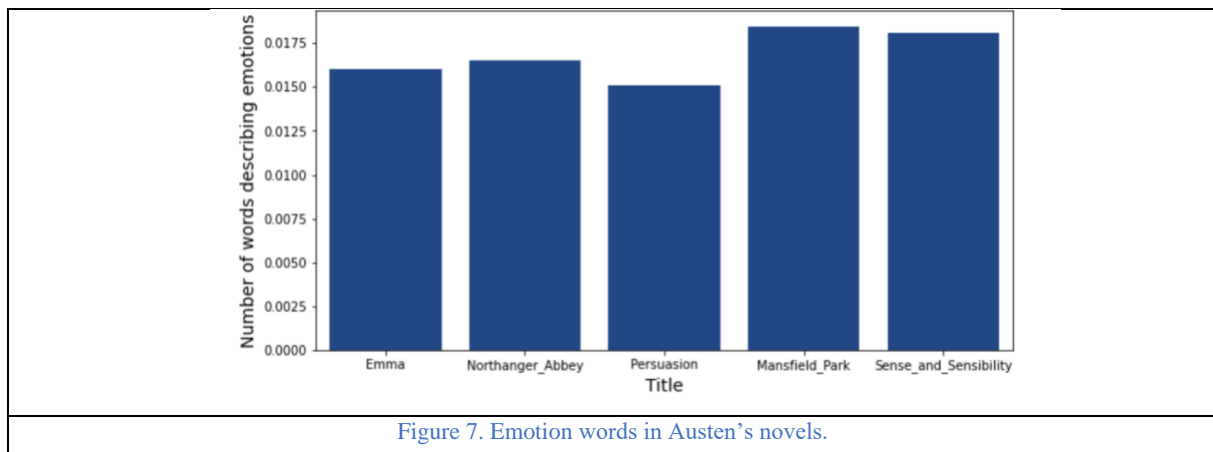
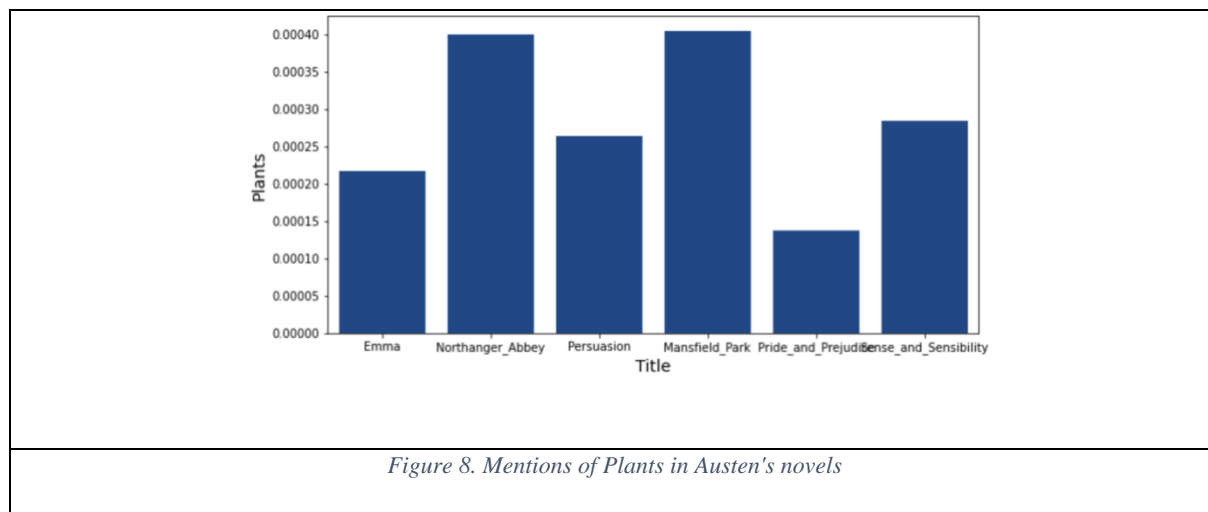
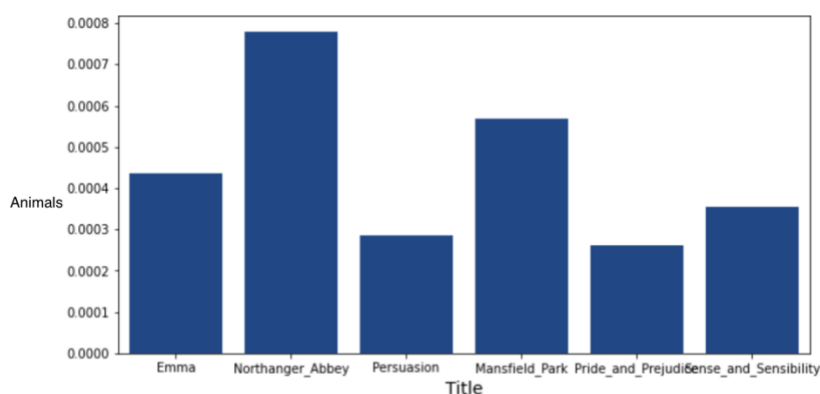


Figure 7. Emotion words in Austen's novels.



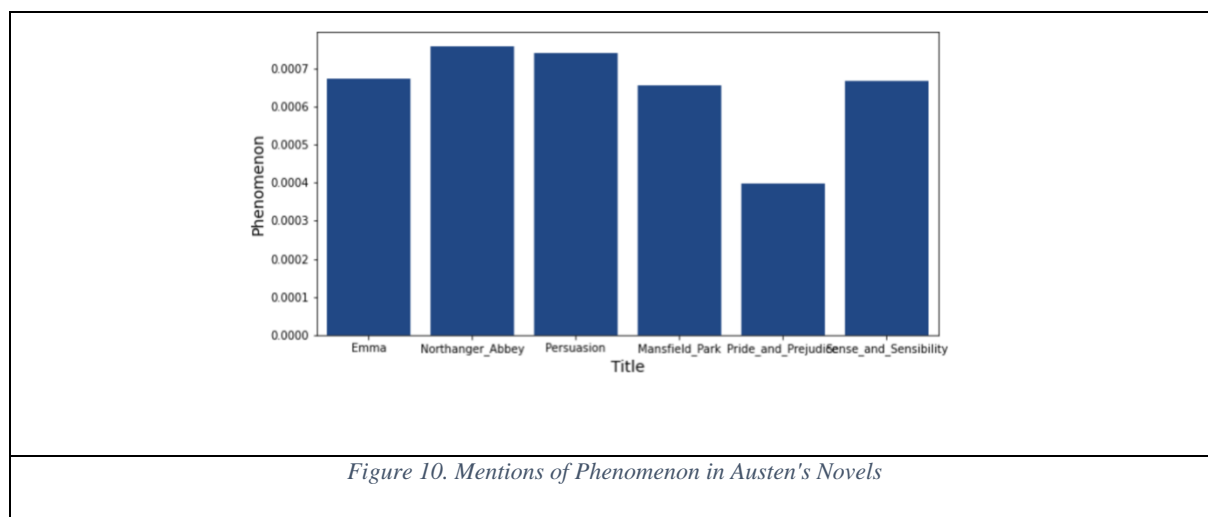
Moreover, the words related to plants reveal how Austen skilfully employs nature imagery to enhance her narratives. Each novel features unique plant-related words, from the ‘blossoms’ of *Emma* to the ‘cedar’ and ‘hyacinths’ of *Northanger Abbey*. In *Mansfield Park*, for example, the description of the beautiful gardens and grounds of the estate plays a role in illustrating the stark contrast between the refined exterior and the moral decay hidden within. The use of the natural world to reflect the characters’ inner worlds is a subtle but effective example of how Austen incorporated nature into her storytelling. I also tested against the frequency of mentions for plants across this novel, and interestingly, *Northanger Abbey* and *Mansfield Park* scored highest compared to the other novels. In contrast, as can be seen in figure 8, *Pride and Prejudice* scores lowest.



The presence of animal-related words in Austen’s novels is another fascinating aspect of her storytelling. Whether it is the cows in *Emma* or the hounds in *Northanger Abbey*, animals are used symbolically and metaphorically, contributing to the narrative’s depth, and meaning. These animal references often reflect characters’ traits and relationships, and their

significance is a subject of intrigue for literary scholars. For example, the mention of cows in *Emma* may symbolise the simplicity and pastoral nature of the countryside, which is contrasted with the more complex and urban social interactions in the novel. In *Northanger Abbey*, hounds may symbolise the pursuit of mysteries and truths, paralleling Catherine Morland's quest to uncover secrets. As can be seen in figure 9, *Northanger Abbey*, scores highest by far compared to the other novels in animal references. *Persuasion* and *Pride and Prejudice* are almost equal in the lowest mentions.

The group of words related to cognition provides further insight into Austen's exploration of characters' thoughts, decisions, and intellectual processes. Her characters ponder, analyse, and make judgements, and their cognition is a crucial element in the unfolding of the plot. The characters in her novels often engage in witty and clever conversations that showcase their cognitive abilities. For instance, in *Emma*, the main character, Emma Woodhouse, is portrayed as an intelligent but sometimes misguided matchmaker. Her attempts to pair people up, while well-intentioned, lead to humorous misunderstandings and serve as examples of cognitive processes like decision-making and social perception. Austen's use of cognitive language reveals her character's thought patterns, enabling readers to connect with their intellectual and emotional journey.



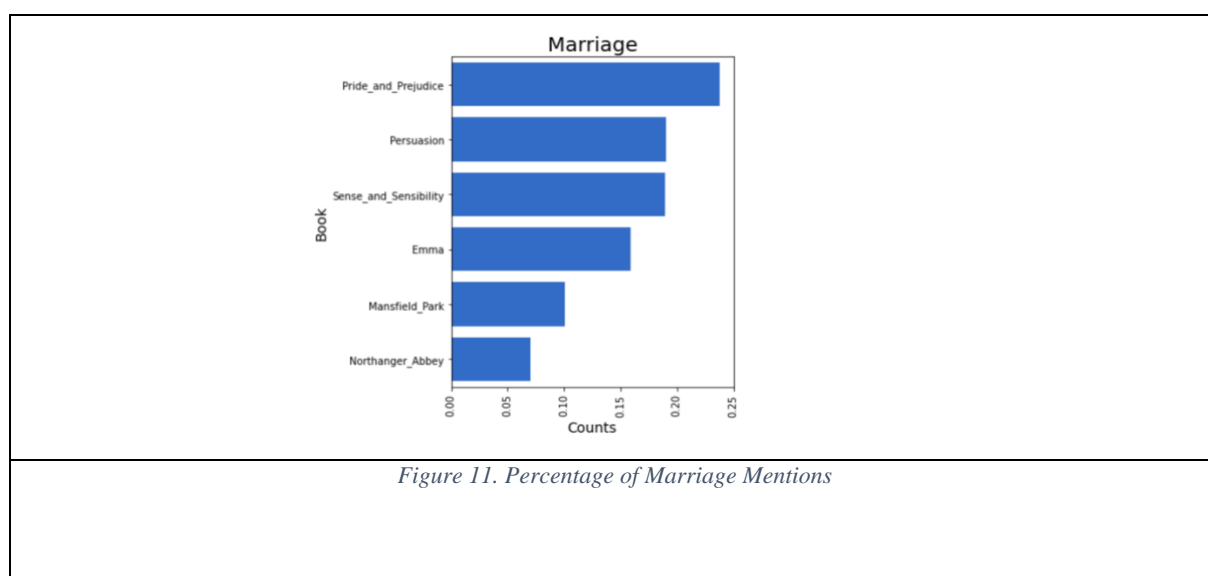
As for natural or supernatural phenomena, in *Northanger Abbey*, Jane Austen humorously parodies the gothic novel genre by exploring the phenomenon of overactive imagination. The protagonist, Catherine Morland, is fascinated by gothic novels and often lets her imagination run wild, creating supernatural phenomena in her mind. This is why the computer flagged words like 'embers', 'tempest', and 'fading light' as unique words related

to phenomena in this novel. Austen’s use of these phenomena serves to comment on the impact of literature and the imagination on individuals in her time. Most of the Austen’s novels are close in range in mentions of phenomena, except for *Pride and Prejudice*, as can be seen in figure 10.

As for the feasibility of computationally analysing metaphors and irony, it remains a complex challenge due to the abstract and context-dependent nature of these literary devices. Metaphors and irony often rely on the subtleties of language and cultural context, making it difficult for automated systems to accurately detect and interpret them. The nuances, tone, and context that Austen employs in her novels to convey metaphors and irony require human intuition and cultural awareness.

### 3.8. Literary Themes

Jane Austen’s novels are celebrated for their insightful exploration of various themes, including marriage and social critique. Through computational analysis, we can gain a deeper understanding of how these themes are woven into her works, while also uncovering new gender-specific word patterns in her character portrayals. These themes were investigated via lexicons: curated lists of words which are taken to refer to a consistent semantic domain. The marriage lexicon, for example, compiled synonyms related to marriage and quantified their occurrences within Austen’s corpus. Additionally, it counted occurrences of marriage-related words. Similar to the marriage lexicon, the social critique lexicon contains words associated with social critique, such as ‘class’, ‘norms’, ‘society’, ‘norm’, and ‘tradition’.



Marriage holds a central position in Jane Austen's novels, reflecting the societal norms and values of her time. A computational analysis of the use of marriage-related words in her works provides fascinating insights into the importance and nuances of this theme across her novels. Among her novels, *Pride and Prejudice* stands out with the highest frequency of marriage-related terms, with approximately 0.235% of the text dedicated to these words, as can be seen in figure 13. This aligns perfectly with the novel's focus on marriage as a significant life event. In *Pride and Prejudice*, Austen delves into the intricacies of matrimony, and words like 'marriage', 'married', 'husband', 'wife', and 'matrimony' are used extensively. These words reflect the novel's exploration of societal pressures and individual desires regarding marriage. In *Persuasion* and *Sense and Sensibility*, the use of marriage-related words remains significant, with about 0.19% of the text dedicated to this theme in both novels. These novels examine the role of marriage in the lives of the characters, shedding light on the societal expectations and financial considerations that often influenced matrimonial decision.

*Emma* and *Mansfield Park* also prominently feature discussions about marriage. While the percentage of marriage-related words is slightly lower in these novels, they still address this theme effectively. *Emma* explores the dynamics of matchmaking, while *Mansfield Park* delves into moral consideration in choosing a spouse. In *Northanger Abbey*, the percentage of marriage-related words is the lowest among Austen's novels, at around 0.07%. This novel takes a more playful approach to the theme, often satirising the conventions of courtship and marriage in the society of the time. Austen's use of synonyms for marriage-related words, such as 'marry', 'union', 'tie', 'espouse', and others, illustrates the diverse ways she approached the topic, providing readers and scholars with a comprehensive view of how marriage was perceived and negotiated during the Regency era.

In addition to marriage, social critique is another recurring theme in Jane Austen's works. Through an analysis of words related to societal norms, class, tradition, and social expectations, we can discern the subtle yet powerful critique she embedded in her narratives.

Among her novels, *Pride and Prejudice* and *Sense and Sensibility* contain a high density of words connected to social critique, with approximately 0.132% and 0.125% of the text dedicated to these themes respectively, as can be seen in figure 15 and 16. Austen's razor-sharp commentary on societal norms, class distinctions, and propriety is vividly portrayed through these words. Her critique serves as a mirror to the societal conventions of her time. *Emma* also offers a substantial portrayal of social critique, with around 0.118% of

the text dedicated to these themes. The novel explores issues of class and societal expectations, showcasing how Austen's characters navigate these challenges while maintaining their individuality. In *Persuasion*, Austen employs words related to social critique to comment on issues such as approbation and respectability. Her critique is delivered with subtlety and wit, offering readers and scholars a lens through which to analyse the societal pressures her characters face.

*Mansfield Park* and *Northanger Abbey* contain slightly lower percentages of words related to social critique, yet they remain essential to the novels' narratives. *Mansfield Park* critiques the moral decay hidden beneath refined exteriors, while *Northanger Abbey* satirises the conventions of gothic novels and the impact of literature on individuals' imaginations.

Apart from thematic analysis, we can use computational methods to explore gender-specific word patterns within Jane Austen's novels. The words spoken by individual characters can be found using BookNLP. The package collects quotes from Austen's works, along with the character name and gender of the speaker. These quotes were linked to specific words, allowing for analysis of word usage within different contexts. This character study delves into unique words for and by female and male characters. In *Emma*, we observe that words like 'husband', 'fancy', 'allowed', and 'girl' are primarily associated with female characters, reflecting their social roles and concerns. In contrast, male characters are often described using words like 'man', 'uncle', 'asking', and 'concerns', showcasing their distinct attributes and responsibilities. *Northanger Abbey* presents a gendered pattern with words such as 'daughters', 'woman', 'desire', 'learned', and 'charge' being more commonly associated with female characters. On the other hand, male characters are depicted through words like 'see', 'horse', 'like', 'delighted', and 'come'. *Persuasion* continues this pattern, with words like 'imagine', 'woman', 'want', 'nerves', and 'staying' mainly used for female characters, while words like 'man', 'wife', 'marriage', 'fortune' and 'proposed' are more frequently associated with male characters.

In *Mansfield Park*, words like 'faults', 'ways', 'inclination', 'comforts', and 'engagement' are employed to describe female characters, while 'man', 'trust', 'hopes', 'dress', and 'means' are used for male characters, highlighting their roles and concerns. *Pride and Prejudice* showcases the gendered language patterns with words like 'ladyship', 'spirits', 'astonishments', 'longed', and 'dislike' primarily linked to female characters. In contrast, male characters are portrayed through words like 'pride', 'lips', 'agreeable', 'fellow', 'applied', 'rejoice', and 'library'. In *Sense and Sensibility*, words like 'husband', 'thoughts', 'determined', 'wonder', and 'help' are predominantly used for female characters. Male

characters are characterised by words like ‘man’, ‘coming’, ‘merits’, ‘embarrassment’, and ‘constancy’. These gendered word patterns provide valuable insights into the distinct roles, concerns, and societal expectations placed upon male and female characters in Austen’s novels.

The analysis of themes and gendered language patterns in Jane Austen’s novels enhances the reading experience and offers valuable material for literary scholars. Readers gain a deeper appreciation of the intricate thematic layers in her works, from the complexities of marriage to her keen societal critique. Furthermore, understanding gendered language patterns provides insights into the nuanced portrayal of male and female characters, shedding light on their societal roles and individual struggles. This knowledge contributes to a more comprehensive understanding of Austen’s character development and the gender dynamics of her time. Incorporating computational methods in literary analysis enables scholars to uncover subtle patterns and themes that might be easily overlooked through traditional close reading. It complements traditional literary analysis by providing quantitative data to support interpretations and offers a new dimension to the study of classic literature.

### *3.9. Conclusion*

The application of computational methods to literary analysis enhances our understanding of literature in several ways. First, it allows us to identify and analyse patterns and themes that might not be immediately evident through traditional close reading. Computational tools can process vast amounts of text, identifying word frequencies, relationships, and unique words, as demonstrated in this analysis of Austen’s novels. Secondly, computational methods provide a quantitative dimension to literary analysis. They can measure and compare word frequencies, sentiment analysis, and other linguistic features across multiple texts, offering a data-driven approach to understanding an author’s style and thematic choices.

The implications of this computational study for literary scholars and professional readers are significant. While technology has advanced in the analysis of text, there are still aspects of literary analysis that require human interpretation and cultural context. The results from this study provide a foundation for deeper exploration of Austen’s works and offer valuable insights into the emotional, natural, and cognitive dimensions of her storytelling. Additionally, the limitations in detecting metaphors and irony computationally serve as a reminder of the irreplaceable role of human intuition and cultural understanding in the

literature. Literary scholars and readers can use these computational findings as a springboard for their analyses, enhancing their appreciation in Austen's timeless narratives.

Certainly, computational research offers a valuable complement to traditional close reading methods in the study of literature. While close reading relies on the meticulous examination of texts, often focusing on themes, characters, and stylistic elements, computational analysis enables researchers to uncover patterns and insights that might be challenging to discern through close reading alone.

Computational tools provide the capacity to analyse extensive bodies of texts efficiently. Jane Austen's works, while individually rich in content, are part of a broader literary corpus. Computational methods can swiftly process and compare her writings, allowing researchers to track changes in her linguistic style across multiple novels. Close reading would require a significant amount of time and effort to achieve this comprehensively. Additionally, computational analysis can identify subtle linguistic shifts. In the case of Jane Austen, the examination of modal verbs and sentence lengths revealed nuanced differences in her writing style across her novels. While a close reader might sense these variations, computational tools can quantify and visualise these trends, making them more apparent.

Moreover, certain findings, like the type-token ratio, provide quantifiable measures to Austen's linguistic diversity compared to her contemporaries. Such quantitative insights, which are challenging to derive from close reading, offer a clearer understanding of how Austen's writing style differed from other authors of her time.

In short, computational research serves as a powerful partner to close reading in literary analysis. It allows scholars to process large datasets efficiently, uncover subtle linguistic shifts, and quantify features in an author's style that could be challenging to extract solely through close reading. By blending these two methods, researchers can gain a deeper understanding of an author's work, like Jane Austen's evolution as a writer and comparing that to her contemporaries, enriching the field of literary analysis.

## Chapter 4 – Conclusion

### 4.1. Introduction

In the course of this research, I have carried out a multifaceted exploration of Jane Austen's literary world. In the first chapter I examined the enduring appeal of Austen's works across centuries, discussing the evolution of reading practices, and introducing the pivotal role of computational methods in contemporary literary analysis. In the second chapter I delved into the dynamics of reading Austen in the 'print age'. This exploration illuminated the intricate relationship between the societal contexts in which Austen's books were consumed, and the diverse interpretations of her novels. The pages of printed books became conduits for cultural dialogues and aesthetic choices, shaping the reader's experience in a unique manner. The print age witnessed an ever-evolving relationship between Austen, her readers, and the texts themselves.

### 4.2. Discussion

Chapter three, which functioned as the core of this research, moved into the digital age, exploring the contemporary reading experience characterised by computational literary analysis. Through this case study, I delved into a diverse array of computational methods to unveil the intricate layers of Austen's narratives. It dissected her linguistic style, probed the depths of her characters' emotions, and unearthed themes embedded in her prose. With the aid of computational tools, I examined sentence lengths, the usage of modal verbs, syntactical structures, alliteration, internal rhyme, emotions, and thematic elements across Austen's novels. The results of the computational research offered several perspectives on the literary genius of Jane Austen. It established the stability in sentence lengths throughout her literary work, the potential thematic depth in *Sense and Sensibility*, and the distinctive linguistic structure that sets her apart from her contemporaries. The case study unveiled the musicality of her prose through alliteration and internal rhyme, shedding light on the subtle poetic elements of her narratives. Moreover, by identifying gender-specific language patterns, I discerned the societal roles and expectations of male and female characters in her novels.

Through this thesis, I have discovered critical insights, inspiring a deeper appreciation for the nuances of Austen's craft. This research, however, is far from complete. In this concluding chapter, I will reflect on the relevance of the case study, offer concrete answers to the research questions I posed at the opening chapter, ponder the future of reading with computational methods, and acknowledge the potential avenues for future research while recognising the limitations of this thesis project. In the following sections I hope to connect

all that was discussed in this research and underscore the significance of my findings in the broader context of literary analysis and the enduring allure of Jane Austen's novels in the digital age.

In chapter 2.6, we discussed Austen's use of diction based on closed reading methods conducted by different literary scholars. The analysis of Jane Austen's narrative craft by Keymer and MacMahon highlights her skillful manipulation of language to evoke specific emotions and themes in her novels<sup>182</sup>. In *Northanger Abbey*, Austen playfully incorporates Gothic elements, which, as is argued by MacMahon, coincides with linguistic shifts.<sup>183</sup> Moreover, Margaret Anne Doody examined the symbolic significance of language in *Emma*, illustrating how seemingly mundane actions, like Mr. Knightley displaying mud-free shoes, reflect deeper societal norms and character dynamics.<sup>184</sup> Austen's linguistic evolution, as noted by Doody and Todd, showcases her ability to infuse vitality into her mature works while retaining thematic depth.<sup>185</sup> Doody suggests a tension between spirited language and hypocrisy, while Todd emphasises Austen's diverse creative journey, questioning the linearity of her stylistic progression.<sup>186</sup>

Chapter 3.4 analyses Austen's diction and vocabulary across her six main novels, *Emma*, *Northanger Abbey*, *Persuasion*, *Mansfield Park*, *Pride and Prejudice* and *Sense and Sensibility*. The analysis highlights recurring thematic and stylistic elements in Austen's writing, such as the prevalence of words like 'marriage', and 'love', reflecting her focus on social conventions and romantic relationships. Austen's unique word choices, predominantly names and familial terms, underscore her attention to character development. The study also examines the usage of the word 'quite', challenging previous scholarly assertions regarding its gender-specific usage in Austen's works.<sup>187</sup> Additionally, the Type-Token ratio analysis reveals *Pride and Prejudice*'s richer vocabulary, suggesting a correlation between lexical variety and novel complexity. Computational tools offer insights into Austen's linguistic evolution and thematic consistency, facilitating a systematic understanding of her narrative techniques and literary contributions.

In chapter 2.6 the main focus was on literary analysis, emphasizing thematic exploration and narrative techniques in Jane Austen's works, such as her use of diction and vocabulary. It delves into the interpretative aspect, drawing conclusions about Austen's

---

<sup>182</sup> Keymer, p. 32. and MacMahon p. 676.

<sup>183</sup> MacMahon, p. 676.

<sup>184</sup> Doody, p. 85.

<sup>185</sup> Doody, p. 86; Todd, p. 87.

<sup>186</sup> *ibid.*

<sup>187</sup> Burrows, p. 64.

thematic concerns and stylistic evolution. Conversely, chapter 3.4. centres on computational analysis, employing quantitative methods to examine Austen's language patterns and lexical richness across her novels. While 2.6. offers qualitative insights into Austen's literary craft, 3.4. provides quantitative data, enhancing understanding through systematic analysis. Together they offer a comprehensive view on Austen's literary prowess from both qualitative and quantitative perspectives.

In chapter 3.5, I analysed Austen's characterisations based on her language choices. Austen employs her diction strategically, revealing the complexities of her characters and the themes they embody. Characters serve as lenses through which readers perceive the dissonance between words and intentions. Austen's characterisations, showcased in works like *Sense and Sensibility* and *Mansfield Park*, pit rounded personas against flat ones, offering nuanced social critiques. Characters like Emma Woodhouse or Elinor Dashwood undergo profound journeys, reflecting Austen's exploration of pride, humility, and resilience. Departing from traditional narratives in *Lady Susan*, Austen challenges societal norms and engages readers in moral contemplation through narrative gaps. Austen's meticulous manuscript revisions and personal experiences further enrich our understanding of her evolving authorial identity. In short, Austen's characters are timeless embodiments of human emotion and societal commentary, captivating readers through generations.

In chapter 3.5, analysing Jane Austen's linguistic structure and comparing her to her contemporaries offers valuable insights into her evolving writing style. Austen's sentence lengths, while generally stable, exhibit variation, with *Sense and Sensibility* featuring longer sentences, possibly indicating deeper thematic exploration. Compared to her peers, Austen tends to use longer sentences, diverging from contemporaries like William Makepeace Thackeray. The use of modal verbs in *Pride and Prejudice* suggests a nuanced portrayal of her characters' inner thoughts, while overall, Austen's modal verb usage aligns with her contemporaries, except for Anthony Trollope, who exceeds her usage. Unique syntactical patterns, especially in *Mansfield Park*, contribute to specific tones or atmospheres. Despite deviations, Austen's novels maintain a coherent stylistic approach, emphasising character development and interpersonal interactions. This syntactical richness underscores Austen's narrative complexity, enhancing both scholarly and general readers' appreciation of her technical prowess and enduring appeal.

Section 3.5 highlighted Austen's skilful portrayal of characters as vibrant personas that transcend the pages, embodying a multitude of traits and qualities, exemplified through examples from her novels. Chapter 3.5. in contrast analyses more quantitative measures, such

as sentence lengths, usage of modal verbs, adjectives, adverbs, and syntactical categories across Austen's works. Although vastly different approaches, the results of both studies confirm that both from a qualitative and quantitative approach, Austen's characterisations confirm her literary genius.

As discussed in chapter 2.7, Jane Austen's narratives are richly embellished by a variety of literary devices, enhancing their depth and complexity. These devices serve as a bridge between the author's intentions and the reader's perception, imbuing the text with layers of meaning. Franco Moretti's insights highlight how examining these devices can trace the evolution of genres over time.<sup>188</sup> Austen's strategic use of irony, imagery, and figures of speech in her novels, notably exemplified in *Mansfield Park*, illuminates character traits, and adds depth to the narrative. Central to Austen's narrative genius is her adept manipulation of irony, evident from the outset of her works, as noted in *Pride and Prejudice*. Penny Gay's analysis delves into Austen's evolving narrative maturity and her increasingly robust satire, particularly evident in *Persuasion*.<sup>189</sup> Austen's mastery extends to her portrayal of books as more than objects, weaving them into her characters' lives to reveal their psychological landscapes.

Austen's novels are celebrated for their intricate language and imagery, and through computational tools, chapter 3.7, attempts to offer insights into quantifiable human emotions, nature, cognition, and phenomena. Emotion, a central theme, is explored through a diverse range of sentiments experienced by characters, exemplified by Elizabeth Bennet's journey in *Pride and Prejudice*. Austen employs nature imagery skilfully, using botanical references to enrich storytelling and evoke sensory experiences, as seen in *Mansfield Park*'s contrasting landscapes. Animals are employed symbolically, enhancing narrative depth, and meaning, with *Northanger Abbey* featuring prominently in animal references. Words related to cognition provide further understanding of characters' thoughts and decisions, as exemplified by Emma Woodhouse's matchmaking endeavors in *Emma*. Austen humorously explores the phenomenon of overactive imagination in *Northanger Abbey*, reflecting on the impact of literature on individuals. Computational analysis of metaphors and irony remains challenging due to their abstract and context-dependent nature, relying on human intuition and cultural awareness for accurate interpretation.

Both chapter 2.7 and 3.7. complement each other by offering different lenses through which to understand Austen's writing. The qualitative analysis in 2.7. sheds light on the

---

<sup>188</sup> Moretti, p. 47.

<sup>189</sup> Gay, p. 56.

artistic and thematic aspects of Austen's works, while the computational analysis in 3.7. provides clear evidence of recurring motifs, themes, and linguistic patterns. It is however evident that computational research is very limited regarding literary devices. Austen's renown is partly based on her use of irony specifically, which is not quantifiable for computational research. Therefore, in this case specifically, it is close reading methods that provide the most thorough understanding of literary devices and their usage. Although this computational study could contribute some, it is not very significant on the scale in Austen's writing.

In exploring Jane Austen's literary domain in chapter 2.4, one encounters a rich tapestry of themes. Margaret Anne Doody and Beth Lau offer contrasting perspectives on Austen's societal commentary. Doody posits Austen's critique as dissecting societal intricacies rather than outright rejection, challenging traditional interpretations.<sup>190</sup> Lau's view on incestuous themes prompts reflection on Austen's narrative depths.<sup>191</sup> The interplay of themes defies categorisation, transcending simplistic boundaries to offer nuanced societal critiques and explore individual agency. Jocelyn Harris and Penny Gay delve into Austen's narrative finesse, examining the interplay of tones, characterizations, and comedic elements.<sup>192</sup> Gay's insights on desire, societal norms, and character dynamics deepen our understanding of Austen's world. Gay also highlights Austen's adeptness at navigating genres, evident in the poetic essence of *Persuasion*.<sup>193</sup> Together, these perspectives reveal Austen's works as multifaceted narratives reflecting timeless themes, societal commentary, and individual agency.

In chapter 3.8, marriage, a central theme, is dissected across her works, with *Pride and Prejudice* prominently featuring discussions, reflecting societal pressures and individual desires. Social critique is another recurring theme, vividly portrayed in *Pride and Prejudice* and *Sense and Sensibility*, offering a mirror to societal conventions. Gender-specific word patterns reveal distinct roles and concerns of male and female characters, with *Emma* and *Northanger Abbey* showcasing unique associations. This analysis enhances the reading experience, providing insights into Austen's thematic depth and character development. By incorporating computational methods, scholars uncover subtle patterns, complementing traditional literary analysis and offering a quantitative dimension to the study of classic literature.

---

<sup>190</sup> Doody, p. 74.

<sup>191</sup> Lau, p. 259.

<sup>192</sup> Harris, p. 53.

<sup>193</sup> Gay, p. 61.

In 2.4, scholars analyze Austen's thematic exploration and narrative finesse, examining themes such as societal critique, individual agency, and the interplay of genres in her novels. Margaret Anne Doody and Beth Lau present contrasting views on Austen's societal commentary, while Jocelyn Harris and Penny Gay delve into the nuances of Austen's narrative techniques and characterizations. This analysis emphasizes the complexity and depth of Austen's works, highlighting her ability to weave timeless themes and societal commentary into captivating narratives. On the other hand, 3.8 focuses on computational analysis to explore themes of marriage, social critique, and gender-specific word patterns in Austen's novels. Through quantitative methods, scholars uncover the frequency and significance of marriage-related words, societal critique, and gendered language patterns across Austen's works. This approach provides empirical evidence to support interpretations of Austen's thematic exploration and character portrayals, offering a quantitative dimension to complement traditional literary analysis.

#### *4.3. Limitations and Future Research*

With this thesis being relatively small, there are naturally a few limitations or neglected aspects to be found. This most importantly relates to contextual understanding, computational methods, while powerful, often struggle with contextual understanding. Metaphors, irony, cultural references, and authorial intent can be challenging for algorithms to interpret accurately. I have attempted to function as the human bridge, interpreting the results from my computational analysis. However, the scope of this research limits the aspects I could focus on, and there are possibilities to interpret the results in a broader context than I have been able to do here. Additionally, it is very probable that there were biases in the computational tools, where things were not flagged as a variation of a word where a 'human' would recognise that variation. This could impact the results of the analysis. Although I filtered through the results, it is possible that I missed elements or biases that were not obvious. This links to the complexity of literary devices. Literary works, such as Jane Austen's, often employ complex devices that are difficult for computational methods to identify and analyse. These include allegory, unreliable narration, multi-layered symbolism, metaphors, and irony. Future research should seek innovative ways to address these challenges and enhance the capabilities of computational analysis. Another aspect that, due to the scope of this project, I neglected is long form reading behaviour. While this study offered insights into linguistic aspects of Jane Austen's works, and hinted at consequences for readers, it did not specifically address long-form reading behaviour and the cognitive engagement of readers.

To enhance and continue the relationship between literary studies and computational methods, there are a few aspects that could inspire further study. Naturally, while this research focused on Jane Austen and her contemporaries, future studies can extend comparative authorship analysis to different literary periods, genres, or authors. By examining how various authors' linguistic styles and thematic preferences evolve across their works, researchers can uncover unique patterns and trends, shedding light on the broader realm of literature. Sentiment analysis, as applied in this study, can be extended to contemporary literature. By analysing modern literary works, researchers can explore how authors of today use language and sentiment to convey complex narratives and emotional depth. Additionally, the results from such a study could be put in comparison to classic authors such as Jane Austen, Charles Dickens, Shakespeare. This could quantify and clearly show the influences that such authors have on today's writers, forming the basis for more longitudinal or diachronic analyses.

In conclusion, the future of computational methods in literary analysis is a journey of continuous exploration, innovation, and refinement. While our research into reading Jane Austen in the digital age has provided valuable insights, it is only a starting point of numerous other possibilities. Future research should strive to overcome these challenges, ensuring that the digital age enriches our understanding of classic and contemporary literature while preserving the depth and nuances that make literature a timeless and profound art form.

As evidenced by the results of this thesis, computational tools play a pivotal role in modern literary analysis. They offer a quantitative lens that can complement traditional qualitative approaches. For Austen's works, computational tools can analyse vast quantities of text, identifying linguistic patterns, unique words, and language structures. These tools enhance the accuracy and efficiency of studying an author's diction. They also allow for large-scale text comparisons, helping researchers understand how an author's language evolves over time, and although not analysed in this study, could be an area of study in which these tools are employed. Computational analysis enables the study of shifts in vocabulary and style, providing valuable insights into an author's artistic growth. Furthermore, these tools facilitate the identification of linguistic markers and help reveal recurring themes and sentiments across an author's canon. In Austen's case, this can help to highlight her consistent exploration of love, marriage, and social interactions.

Computational tools can also influence the reading experience for literary scholars. They provide a structured approach to text analysis, fostering a more systematic understanding of an author's linguistic choices. This can aid scholars in drawing nuanced

conclusions about an author's thematic concerns, stylistic evolution, and unique literary contributions. Additionally, after conducting a computational study on an author, the next time the scholar reads the works of the specific author, it might make them notice other elements that they would not have before.

## Bibliography

- Brooke, Julian, Adam Hammond, and Graeme Hirst, 'Using Models of Lexical Style to Quantify Free Indirect Discourse in Modernist Fiction' *Digital Scholarship in the Humanities* vol 32. Issue 2. 2017, pp. 234-50.
- Brooks, Cleanth, *The Well-Wrought Urn: Studies in the Structure of Poetry*, (London: Dobson, 1968).
- Burrows, J. F., *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*, (New York: Oxford University Press, 1987).
- Caddy, Scott, 'Fan Media and Transmedia: Jane Austen in the Digital Age.', in *Fan Phenomena: Jane Austen*, ed. by Gabrielle Malcolm (Bristol: Intellect Books, 2015), pp. 42-51.
- DeForest, Mary, and Eric Johnson., 'The Density of Latinate Words in the Speeches of Jane Austen's Characters' *Literary and Linguistic Computing* vol 16. Issue 4. 2001, pp. 389-401.
- Doody, Margaret Anne., 'The Early Short Fiction' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 72-86.
- DuBois, Andrew., 'Close Reading: An Introduction' in *Close Reading: The Reader*, ed. By Frank Lentricchia and Andrew DuBois (Durham: Duke University Press, 2003), pp. 1-40.
- Dunning, Ted. 'Accurate Methods for the Statistics of Surprise and Coincidence' *Computational Linguistics*, vol 19 issue 1. 1993, pp. 61-74.
- Edwards, Alexandra. 'In the Regency Alternate Universe: Jane Austen and Fanfiction Culture.', in *Fan Phenomena: Jane Austen* ed. by Gabrielle Malcolm (Bristol: Intellect Books, 2015), pp. 10-19.
- Eve, Martin Paul., *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas*, (Stanford: Stanford University Press, 2019).
- Fergus, Jan., 'The Professional Woman Writer' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 1-20.

- Gay, Penny., 'Emma and Persuasion' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 55-71.
- Grundy, Isobel., 'Jane Austen and Literary Traditions' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 192-214.
- Harris, Jocelyn., 'Pride and Prejudice and Mansfield Park' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 39-54.
- Keymer, Thomas., 'Northanger Abbey and Sense and Sensibility' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 21-38.
- Lau, Beth. 'Placing Jane Austen in the Romantic Period: Self and Solitude in the Works of Austen and the Male Romantic Poets' *European Romantic Review* vol. 15 issue 2. 2004, pp. 255-67.
- Lynch, Deidre. *Janeites: Austen's Disciples and Devotees*. (Princeton: Princeton UP, 2000)
- MacMahon, Barbara., 'Metarepresentation and Decoupling in *Northanger Abbey*: Part 2' *English Studies* vol 90. Issue 6. 2009, pp. 673-94.
- Mann, H.B., and D.R. Whitney. 'On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other' *Ann. Math. Statist.* Vol 18 issue 1. 1947, pp. 50-60.
- Mattingly, W.J. B. 'Introduction to BookNLP'  
<[https://booknlp.pythonhumanities.com/01\\_intro.html](https://booknlp.pythonhumanities.com/01_intro.html)> (17 January 2024),
- Moretti, Franco., *Distant Reading* (London: Verso, 2013).
- Oxford Dictionary of Literary Terms 4<sup>th</sup> ed. 'Alliteration' <<https://www-oxfordreference-com.ezproxy.leidenuniv.nl/display/10.1093/acref/9780198715443.001.0001/acref-9780198715443-e-31?rskey=fhORwt&result=1>> (1 June 2024).
- Oxford Dictionary of Literary Terms 4<sup>th</sup> ed. 'Internal Rhyme' <<https://www-oxfordreference-com.ezproxy.leidenuniv.nl/display/10.1093/acref/9780198715443.001.0001/acref-9780198715443-e-603?rskey=ETBkA0&result=1>> (1 June 2024).
- Paris, Bernard J. *Character and Conflict in Jane Austen's Novels: A Psychological Approach* (Detroit: Wayne State UP, 1978).
- Richards, I.A., 'The Analysis of a Poem' *Principles of Literary Criticism*, fourth edition, (London: Kegan Paul Trench, Trubner & Co LTD, 1930), pp. 114-33.

- Rigney, Anne. 'The Many Dimensions of Literature' in *The Life of Texts: An Introduction to Literary Studies*, ed. by Kiene Brillenburg Wurth, and Anne Rigney. (Amsterdam: Amsterdam UP, 2019), pp. 43-75.
- Todd, Janet., 'Lady Susan, The Watsons, and Sandition' in *The Cambridge Companion to Jane Austen*, ed. by Edward Copeland and Juliet McMaster (Cambridge: Cambridge University Press, 2011), pp. 87-96.
- Verhaar, Peter, *Affordances and Limitations of Algorithmic Criticism*, (unpublished doctoral thesis, 2016)
- Weel, Adriaan, van der., 'Feeding Our Reading Machines: From the Typographic Page to the Docuverse' *Digital Studies* vol 6. 29-01-2017  
<<https://www.digitalstudies.org/articles/10.16995/dscn.15/>> (21 June 2023).
- , 'Salient Features of Digital Textuality' in *Changing Our Textual Minds: Towards a Digital Order of Knowledge* (Manchester: Manchester University Press, 2011), pp. 142-92.
- Wells, Juliette. *Everybody's Jane: Austen in the Popular Imagination* (London: Bloomsbury Publishing, 2012).

*Primary sources/ Austen Corpus:*

- Austen, Jane, *Emma*, (1815) Project Gutenberg, txt-file. Updated March 2018.
- Austen, Jane, *Northanger Abbey*, (1817) Project Gutenberg, txt-file. Updated Jan 2021.
- Austen, Jane, *Pride and Prejudice*, (1813) Project Gutenberg, txt-file. Updated Jan 2021.
- Austen, Jane, *Sense and Sensibility*, (1811) Project Gutenberg, txt-file. Updated Dec 2020.
- Austen, Jane., *Mansfield Park*, (1814) Project Gutenberg, txt-file. Updated March 2018.
- Austen, Jane., *Persuasion*, (1817) Project Gutenberg, txt-file. Updated Feb 2015.

## Appendices

### *Appendix A*

<http://www.gutenberg.org/files/5826/5826-8.txt>  
<http://www.gutenberg.org/files/6042/6042.txt>  
<http://www.gutenberg.org/files/6053/6053.txt>  
<http://www.gutenberg.org/files/6346/6346-0.txt>  
<http://www.gutenberg.org/files/6457/6457.txt>  
<http://www.gutenberg.org/files/7146/7146-8.txt>  
<http://www.gutenberg.org/files/7152/7152-0.txt>  
<http://www.gutenberg.org/files/29125/29125.txt>  
<http://www.gutenberg.org/files/37437/37437.txt>  
<http://www.gutenberg.org/files/37438/37438.txt>  
<http://www.gutenberg.org/files/37439/37439-8.txt>  
<http://www.gutenberg.org/files/37440/37440.txt>  
<http://www.gutenberg.org/files/37441/37441-8.txt>  
<http://www.gutenberg.org/files/40619/40619-8.txt>  
<http://www.gutenberg.org/files/57472/57472-0.txt>  
<http://www.gutenberg.org/files/58079/58079-0.txt>  
<http://www.gutenberg.org/files/50054/50054-8.txt>  
<http://www.gutenberg.org/files/15034/15034-8.txt>  
<http://www.gutenberg.org/files/19595/19595.txt>  
<http://www.gutenberg.org/files/30524/30524-8.txt>  
<http://www.gutenberg.org/files/31697/31697-8.txt>  
<http://www.gutenberg.org/files/31879/31879-8.txt>  
<http://www.gutenberg.org/files/36256/36256-8.txt>  
<http://www.gutenberg.org/files/36257/36257-8.txt>  
<http://www.gutenberg.org/files/37774/37774.txt>  
<http://www.gutenberg.org/files/41646/41646.txt>  
<http://www.gutenberg.org/files/84/84-0.txt>  
<http://www.gutenberg.org/files/4695/4695.txt>  
<http://www.gutenberg.org/files/6447/6447.txt>  
<http://www.gutenberg.org/files/6542/6542-readme.txt>  
<http://www.gutenberg.org/files/15238/15238-8.txt>  
<http://www.gutenberg.org/files/18247/18247.txt>  
<http://www.gutenberg.org/files/20038/20038-readme.txt>  
<http://www.gutenberg.org/files/41445/41445.txt>  
<http://www.gutenberg.org/files/42324/42324-8.txt>  
<http://www.gutenberg.org/files/56665/56665-0.txt>  
<http://www.gutenberg.org/files/768/768.txt>  
<http://www.gutenberg.org/files/5196/5196-0.txt>  
<http://www.gutenberg.org/files/11936/11936.txt>  
<http://www.gutenberg.org/files/19197/19197.txt>  
<http://www.gutenberg.org/files/767/767.txt>  
<http://www.gutenberg.org/files/969/969-0.txt>  
<http://www.gutenberg.org/files/6124/6124.txt>  
<http://www.gutenberg.org/files/9296/9296.txt>  
<http://www.gutenberg.org/files/9798/9798.txt>  
<http://www.gutenberg.org/files/9881/9881.txt>

<http://www.gutenberg.org/files/10462/10462-8.txt>  
<http://www.gutenberg.org/files/10799/10799.txt>  
<http://www.gutenberg.org/files/11364/11364.txt>  
<http://www.gutenberg.org/files/11889/11889-8.txt>  
<http://www.gutenberg.org/files/12180/12180-8.txt>  
<http://www.gutenberg.org/files/12398/12398-8.txt>  
<http://www.gutenberg.org/files/12958/12958.txt>  
<http://www.gutenberg.org/files/13884/13884-8.txt>  
<http://www.gutenberg.org/files/24860/24860.txt>  
<http://www.gutenberg.org/files/29964/29964.txt>  
<http://www.gutenberg.org/files/1146/1146.txt>  
<http://www.gutenberg.org/files/1147/1147-0.txt>  
<http://www.gutenberg.org/files/5256/5256.txt>  
<http://www.gutenberg.org/files/6095/6095.txt>  
<http://www.gutenberg.org/files/6096/6096.txt>  
<http://www.gutenberg.org/files/6097/6097.txt>  
<http://www.gutenberg.org/files/6098/6098-0.txt>  
<http://www.gutenberg.org/files/6593/6593-0.txt>  
<http://www.gutenberg.org/files/6828/6828.txt>  
<http://www.gutenberg.org/files/9609/9609.txt>  
<http://www.gutenberg.org/files/9611/9611.txt>  
<http://www.gutenberg.org/files/30962/30962.txt>  
<http://www.gutenberg.org/files/37923/37923.txt>  
<http://www.gutenberg.org/files/43520/43520-8.txt>  
<http://www.gutenberg.org/files/58249/58249-0.txt>  
<http://www.gutenberg.org/files/599/599.txt>  
<http://www.gutenberg.org/files/897/897-0.txt>  
<http://www.gutenberg.org/files/1462/1462-0.txt>  
<http://www.gutenberg.org/files/1863/1863.txt>  
<http://www.gutenberg.org/files/1933/1933.txt>  
<http://www.gutenberg.org/files/1935/1935.txt>  
<http://www.gutenberg.org/files/1969/1969-0.txt>  
<http://www.gutenberg.org/files/1985/1985-0.txt>  
<http://www.gutenberg.org/files/1990/1990-0.txt>  
<http://www.gutenberg.org/files/2511/2511-0.txt>  
<http://www.gutenberg.org/files/2608/2608.txt>  
<http://www.gutenberg.org/files/2645/2645.txt>  
<http://www.gutenberg.org/files/2646/2646.txt>  
<http://www.gutenberg.org/files/2648/2648.txt>  
<http://www.gutenberg.org/files/2675/2675-0.txt>  
<http://www.gutenberg.org/files/2686/2686-0.txt>  
<http://www.gutenberg.org/files/2731/2731.txt>  
<http://www.gutenberg.org/files/2732/2732.txt>  
<http://www.gutenberg.org/files/2768/2768-0.txt>  
<http://www.gutenberg.org/files/2796/2796-0.txt>  
<http://www.gutenberg.org/files/2797/2797.txt>  
<http://www.gutenberg.org/files/2823/2823-0.txt>  
<http://www.gutenberg.org/files/2843/2843.txt>

<http://www.gutenberg.org/files/2844/2844.txt>  
<http://www.gutenberg.org/files/2859/2859.txt>  
<http://www.gutenberg.org/files/4558/4558-0.txt>  
<http://www.gutenberg.org/files/7265/7265-0.txt>  
<http://www.gutenberg.org/files/7467/7467.txt>  
<http://www.gutenberg.org/files/8123/8123-0.txt>  
<http://www.gutenberg.org/files/9517/9517-readme.txt>  
<http://www.gutenberg.org/files/9518/9518-readme.txt>  
<http://www.gutenberg.org/files/9519/9519-readme.txt>  
<http://www.gutenberg.org/files/9520/9520-readme.txt>  
<http://www.gutenberg.org/files/9521/9521-readme.txt>  
<http://www.gutenberg.org/files/9522/9522-readme.txt>  
<http://www.gutenberg.org/files/9523/9523-readme.txt>  
<http://www.gutenberg.org/files/9524/9524-readme.txt>  
<http://www.gutenberg.org/files/9525/9525-readme.txt>  
<http://www.gutenberg.org/files/9526/9526-readme.txt>  
<http://www.gutenberg.org/files/9527/9527-readme.txt>  
<http://www.gutenberg.org/files/9528/9528-readme.txt>  
<http://www.gutenberg.org/files/9529/9529-readme.txt>  
<http://www.gutenberg.org/files/9530/9530-readme.txt>  
<http://www.gutenberg.org/files/9531/9531-readme.txt>  
<http://www.gutenberg.org/files/9533/9533-readme.txt>  
<http://www.gutenberg.org/files/9534/9534-readme.txt>  
<http://www.gutenberg.org/files/9535/9535-readme.txt>  
<http://www.gutenberg.org/files/9536/9536-readme.txt>  
<http://www.gutenberg.org/files/9537/9537-readme.txt>  
<http://www.gutenberg.org/files/9538/9538-readme.txt>  
<http://www.gutenberg.org/files/9539/9539-readme.txt>  
<http://www.gutenberg.org/files/9540/9540-readme.txt>  
<http://www.gutenberg.org/files/9541/9541-readme.txt>  
<http://www.gutenberg.org/files/9904/9904-8.txt>  
<http://www.gutenberg.org/files/10558/10558-m/10558-m-readme.txt>  
<http://www.gutenberg.org/files/29363/29363.txt>  
<http://www.gutenberg.org/files/34922/34922.txt>  
<http://www.gutenberg.org/files/42890/42890.txt>  
<http://www.gutenberg.org/files/45554/45554-8.txt>  
<http://www.gutenberg.org/files/45921/45921.txt>  
<http://www.gutenberg.org/files/57966/57966-0.txt>  
<http://www.gutenberg.org/files/619/619-8.txt>  
<http://www.gutenberg.org/files/814/814.txt>  
<http://www.gutenberg.org/files/1865/1865.txt>  
<http://www.gutenberg.org/files/1866/1866-8.txt>  
<http://www.gutenberg.org/files/2158/2158.txt>  
<http://www.gutenberg.org/files/2432/2432.txt>  
<http://www.gutenberg.org/files/2860/2860.txt>  
<http://www.gutenberg.org/files/3045/3045.txt>  
<http://www.gutenberg.org/files/3166/3166.txt>  
<http://www.gutenberg.org/files/3409/3409-8.txt>

<http://www.gutenberg.org/files/3550/3550-0.txt>  
<http://www.gutenberg.org/files/3615/3615-0.txt>  
<http://www.gutenberg.org/files/3616/3616-0.txt>  
<http://www.gutenberg.org/files/3622/3622.txt>  
<http://www.gutenberg.org/files/3699/3699-0.txt>  
<http://www.gutenberg.org/files/3700/3700-0.txt>  
<http://www.gutenberg.org/files/3710/3710-0.txt>  
<http://www.gutenberg.org/files/3711/3711-0.txt>  
<http://www.gutenberg.org/files/3712/3712-0.txt>  
<http://www.gutenberg.org/files/3713/3713-0.txt>  
<http://www.gutenberg.org/files/3716/3716-0.txt>  
<http://www.gutenberg.org/files/3717/3717-0.txt>  
<http://www.gutenberg.org/files/3718/3718-0.txt>  
<http://www.gutenberg.org/files/3719/3719-0.txt>  
<http://www.gutenberg.org/files/3720/3720-0.txt>  
<http://www.gutenberg.org/files/3723/3723-0.txt>  
<http://www.gutenberg.org/files/3724/3724-0.txt>  
<http://www.gutenberg.org/files/3767/3767-0.txt>  
<http://www.gutenberg.org/files/4599/4599.txt>  
<http://www.gutenberg.org/files/4917/4917-8.txt>  
<http://www.gutenberg.org/files/4969/4969.txt>  
<http://www.gutenberg.org/files/5118/5118.txt>  
<http://www.gutenberg.org/files/5140/5140-8.txt>  
<http://www.gutenberg.org/files/5202/5202.txt>  
<http://www.gutenberg.org/files/5231/5231.txt>  
<http://www.gutenberg.org/files/5642/5642.txt>  
<http://www.gutenberg.org/files/5709/5709-8.txt>  
<http://www.gutenberg.org/files/5897/5897-8.txt>  
<http://www.gutenberg.org/files/5978/5978.txt>  
<http://www.gutenberg.org/files/7381/7381-8.txt>  
<http://www.gutenberg.org/files/7481/7481-8.txt>  
<http://www.gutenberg.org/files/8897/8897.txt>  
<http://www.gutenberg.org/files/8945/8945-8.txt>  
<http://www.gutenberg.org/files/11643/11643-8.txt>  
<http://www.gutenberg.org/files/12234/12234.txt>  
<http://www.gutenberg.org/files/15766/15766-8.txt>  
<http://www.gutenberg.org/files/16804/16804-8.txt>  
<http://www.gutenberg.org/files/18000/18000-8.txt>  
<http://www.gutenberg.org/files/18640/18640-8.txt>  
<http://www.gutenberg.org/files/18645/18645.txt>  
<http://www.gutenberg.org/files/19500/19500-8.txt>  
<http://www.gutenberg.org/files/21847/21847.txt>  
<http://www.gutenberg.org/files/22000/22000.txt>  
<http://www.gutenberg.org/files/23000/23000.txt>  
<http://www.gutenberg.org/files/24000/24000-8.txt>  
<http://www.gutenberg.org/files/24103/24103-8.txt>  
<http://www.gutenberg.org/files/25001/25001.txt>  
<http://www.gutenberg.org/files/25579/25579.txt>

<http://www.gutenberg.org/files/26001/26001-8.txt>  
<http://www.gutenberg.org/files/26002/26002.txt>  
<http://www.gutenberg.org/files/26287/26287-readme.txt>  
<http://www.gutenberg.org/files/26541/26541.txt>  
<http://www.gutenberg.org/files/27067/27067.txt>  
<http://www.gutenberg.org/files/27533/27533.txt>  
<http://www.gutenberg.org/files/27712/27712.txt>  
<http://www.gutenberg.org/files/28676/28676.txt>  
[http://www.gutenberg.org/files/28811/28811\\_readme.txt](http://www.gutenberg.org/files/28811/28811_readme.txt)  
<http://www.gutenberg.org/files/29000/29000-8.txt>  
<http://www.gutenberg.org/files/29828/29828.txt>  
<http://www.gutenberg.org/files/30100/30100-8.txt>  
<http://www.gutenberg.org/files/30606/30606-8.txt>  
<http://www.gutenberg.org/files/31274/31274.txt>  
<http://www.gutenberg.org/files/33500/33500.txt>  
<http://www.gutenberg.org/files/34000/34000.txt>  
<http://www.gutenberg.org/files/44992/44992.txt>  
<http://www.gutenberg.org/files/54424/54424-0.txt>  
<http://www.gutenberg.org/files/54783/54783-0.txt>  
<http://www.gutenberg.org/files/55100/55100.txt>  
<http://www.gutenberg.org/files/55147/55147-0.txt>  
<http://www.gutenberg.org/files/55212/55212-0.txt>  
<http://www.gutenberg.org/files/55648/55648-0.txt>  
<http://www.gutenberg.org/files/55665/55665-0.txt>  
<http://www.gutenberg.org/files/55926/55926-0.txt>  
<http://www.gutenberg.org/files/58383/58383-0.txt>  
<http://www.gutenberg.org/files/58558/58558-0.txt>

### *Appendix B*

Folder: <https://github.com/MMvVeenen/austen-in-the-digital-age>

BookNLP: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/BookNLP.ipynb>

Characters: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Characters.ipynb>

Contemporary: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Contemporary.ipynb>

Context: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Context.ipynb>

Creating the Corpus: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Creating%20the%20corpus.ipynb>

Diction compared to contemporary authors: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Diction%20compared%20to%20contemporary%20authors.ipynb>

Lexicon Marriage: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Lexicon%20Marriage.ipynb>

Lexicon Social Critique: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Lexicon%20Social%20Critique.ipynb>

Named entities: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Named%20entities.ipynb>

POS\_TTR\_Austen&Contemporary\_Authors: [https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/POS\\_TTR\\_Austen%26Contemporary\\_Authors.ipynb](https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/POS_TTR_Austen%26Contemporary_Authors.ipynb)

Poetic Elements: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Poetic%20elements.ipynb>

Quotes: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Quotes.ipynb>

Sentence Length\_POS\_TTR: [https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Sentence%20Length\\_POS\\_TTR.ipynb](https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Sentence%20Length_POS_TTR.ipynb)

Supersense Entities: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/Supersense%20Entities.ipynb>

Data:

Data\_contemporaries: [https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/data\\_contemporaries.csv](https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/data_contemporaries.csv)

Lexicon: <https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/lexicon.txt>

Lexicon\_words: [https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/lexicon\\_words.csv](https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/lexicon_words.csv)

Lexicon\_words2: [https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/lexicon\\_words2.csv](https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/lexicon_words2.csv)

Sentences\_containing\_lexicon\_words: [https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/sentences\\_containing\\_lexicon\\_words.txt](https://github.com/MMvVeenen/austen-in-the-digital-age/blob/main/sentences_containing_lexicon_words.txt)