Exploring Vocabulary Patterns through Digital Humanities: A Comparative Tool-Based Study on Arundhati Roy's Novels

Thirumoorthi K.[1], Thamil Selvi V.[2]

[1]Research Scholar, [2]Associate Professor, Department of English, School of Engineering and Technology, Dhanalakshmi Srinivasan University, Tiruchirappalli, Tamilnadu, India.

E-mail: [1]thirumoorthik.phd2023@dsuniversity.ac.in, [2]thamilselviv.set@dsuniversity.ac.in

Abstract

Vocabulary analysis is a valuable tool in literary studies, particularly when examining complex texts with layered meanings. The emergence of Digital Humanities has significantly simplified the process of vocabulary analysis. These digital tools assist researchers in identifying vocabulary, visualising vocabulary patterns, tracking changes across text, checking readability, monitoring repetition and lexical density, and comparing large textual units efficiently. Vocabulary serves as a structural and symbolic thread in literary texts, and Digital Humanities tools provide new avenues for tracing this thread. By bridging the gap between traditional literary methods and digital analysis, there is scope for enhanced understanding of narrative voice and the linguistic fabric of fiction. The deliberate repetition of specific terms or phrases often signals larger concerns within the narrative. This paper focuses on two accessible Digital Humanities tools, Voyant Tools and Online-Utility.org, which need not be installed but can be used free online. It also explores how these tools aid vocabulary analysis for literary interpretation. Further, these platforms enable researchers to explore how certain words recur across different narrative segments or in relation to particular characters or events, offering insights into the symbolic frameworks of texts. The novels The God of Small Things and The Ministry of Utmost Happiness by Arundhati Roy are used to explore vocabulary. It specifies the best tool for better literary support without coding knowledge in navigating the intricacies of vocabulary and achieving deeper textual insights.

Keywords: Digital humanities, Vocabulary analysis, Voyant Tools, Online-Utility.org, Arundhati Roy.

1. Introduction

The advent of Digital Humanities has presented new opportunities for researchers to study language and literature using technological tools. The interdisciplinary field combines computational methods with humanities research and permits greater textual analysis, pattern detection and the visual representation of linguistic patterns that would be hard to distinguish through the conventional close reading. Even with these advances, there remains a significant gap in accessibility; most of the textual analysis tools currently available necessitate a certain level of technical expertise. Understanding coding languages or corpus annotation is very difficult for non-coders, especially those from humanities backgrounds, who might be hindered in their ability to fully participate in digital methodologies. Vocabulary forms the foundation of literary language, determining not merely the tone and style of a story but also the emotional and thematic richness contained within. In literature, a close reading of vocabulary provides a means by which subtleties of meaning, character development, and narrative structure can be better comprehended. From symbolism and repetition to tone and readability, word selection and word arrangement are central to the reader's interpretive journey. Therefore, there is an increasing necessity for user-friendly, browser-based systems that facilitate the democratization of digital analysis without compromising depth or functionality. This research seeks to meet that need by comparing two easily available, freely accessible tools Voyant Tools and Online-Utility.org that support vocabulary analysis without the need to know how to program. These digital tools assist researchers in identifying vocabulary, visualising vocabulary patterns, tracking changes across text, checking readability, monitoring repetition and lexical density, and comparing large textual units efficiently. Both tools are used directly online and permit users to input text for immediate, rich feedback on vocabulary form, richness, lexical diversity, frequency, range, density, and complexity.

Vocabulary forms the foundation of literary language, determining not merely the tone and style of a story but also the emotional and thematic richness contained within. In literature, a close reading of vocabulary provides a means by which subtleties of meaning, development of character, and structure of narrative can be comprehended more effectively. From symbolism and repetition to tone and readability, word selection and word arrangement are central to the reader's interpretive journey. Online-Utility.org Text Analyzer is a freely available, web-based software that provides elementary textual analysis without needing any installation and is well-suited to users without programming skills, especially researchers who

are not familiar with programming environments, thus addressing an important usability gap in Digital Humanities research. One of the key features of this tool is its ability to provide quick lexical statistics such as overall word count, average word length, readability scores, and vocabulary density. These provide a rich understanding of the linguistic style of the author and the overall readability of the text. For example, readability measures like the Flesch Reading Ease score and Gunning Fog Index aid in determining the level of complexity in prose, which is particularly important in works of literature characterized by complex sentence structures or unusual narrative structures. Moreover, the software provides keyword density and stop word and filler phrase identification. It also allows for an easy comparison between two distinct texts through the manual scrutiny of frequency lists and pattern trends, and it is helpful in detecting repetitive thematic vocabulary or stylistic conventions within an author's corpus. Although it has its strengths, the tool does have some limitations. The lack of visual output, limited corpus segmentation support, and inability to track keyword trends directly over time may limit its use for more sophisticated literary modeling. However, its accessibility, speed, and comprehensive lexical statistics make it a useful tool for initial vocabulary mapping, especially in comparative or thematic analysis studies. This paper proceeds by offering a comparative analysis of Voyant Tools and Online-Utility.org, examining their individual affordances and limitations in literary vocabulary analysis, with the aim of evaluating their broader potential for use in Digital Humanities research.

1.1 Problem Statement

Despite the growing integration of Digital Humanities in literary studies, limited attention has been given to how accessible, non-coding tools support vocabulary analysis in complex literary texts. Existing research often overlooks comparative assessments of such tools, particularly in their ability to facilitate vocabulary analysis. This study addresses this gap by examining the effectiveness of Voyant Tools and Online-Utility.org Text Analyzer in interpreting the vocabulary of Arundhati Roy's novels, aiming to support literary scholars who seek depth without programming expertise.

2.   Literature Review

Digital Humanities emerged as a response to the overwhelming data in the digital age, such as the vast amounts of digital information that were previously unmanageable. The potential of Digital Humanities can be understood through Busa's work in the 1950s and 1960s,

which has now become feasible in mere minutes using digital tools. As Kim Lacey notes, "The digital humanities were awesome! We're seeing increased cross-fertilization between areas like writing and science, literature and computers, history and engineering" (Lacey 284). Digital Humanities can be categorized into three waves. The first wave focused on the development of the infrastructure necessary for text analysis. During this phase, qualitative software research was conducted, and elements were extracted from databases and archives. Additionally, studies emphasized both quantitative and corpus analyses (Svensson). The second wave emphasized the digital analysis of e-literature, web mining, and the creation of digital toolkits to study cyberculture and digital culture. The fundamental building blocks of digital culture, computer codes, were explored, and digital humanities tools were designed to uncover hidden patterns. In this wave, both qualitative and interpretive research studies were conducted using digital toolkits. The third wave, which emerged as computational capabilities advanced, shifted focus to a broader exploration of the computational aspects of text. The paradigm of close reading gave way to distant reading (Moretti), highlighting a methodological shift. Liu discusses anomalies found in Digital Humanities tools and analyses, demonstrating the evolution of the field into more complex computational methods (Liu). This era brought significant developments in distant reading, exemplified by the use of Voyant Tools, which allow for the visualization and exploration of data at larger scales.

Aathira et al. discuss the challenges and opportunities that Digital Humanities presents, particularly in language and literature studies. They note that technical expertise can pose challenges for scholars unfamiliar with programming, and ethical considerations regarding data privacy and the interpretation of quantitative results require careful attention. However, the authors emphasize that the opportunities outweigh the challenges, as Digital Humanities democratizes access to knowledge and enhances teaching and learning through engaging interactive tools for students (Aathira et al.). Knight et al. examine the development of FreeTxt, a bilingual corpus-based toolkit designed to facilitate the systematic analysis and visualization of free-text survey and questionnaire data. They highlight the potential for expanding the tool to enable comparisons within datasets, map patterns over time, and integrate external sources such as social media. The authors stress that collaboration between academics and end-users plays a crucial role in solving real-world problems and developing effective tools (Knight et al.). The issue of balancing stability and adaptability in vocabulary analysis is another challenge. The vocabulary must remain stable enough to be historically meaningful, while also adapting to the evolving nature of language. Recent advancements in distributional methods,

such as word embeddings, allow for a more data-driven approach to identifying semantically related words over time (Hamilton et al.). This approach, used in historical scholarship, offers new possibilities for studying the evolution of language. In recent years, vocabulary analysis has become a prominent area of study within Digital Humanities, especially through the application of web-based tools that assist researchers in uncovering linguistic patterns, frequencies, and distributions. Tools such as Voyant Tools and Online-Utility.org Text Analyzer have been instrumental in providing interactive environments for textual analysis. Voyant Tools supports multifaceted visualization such as word clouds, frequency graphs, correlation analysis, and keyword trends that allow for both surface-level and in-depth analysis of language use. Similarly, Online-Utility.org enables straightforward vocabulary density evaluation and lexical structure analysis. These tools are especially valuable for scholars without coding expertise, making vocabulary analysis more accessible and applicable to large literary texts. Their combined application to literary works reveals recurring themes, stylistic choices, and semantic shifts, opening new interpretative avenues for digital literary criticism.

Luhmann and Burghardt analyze the role and position of Digital Humanities in the academic landscape, proposing a model for understanding the relationships between topics, articles, disciplines, and journals. They stress that a multidimensional view of Digital Humanities, incorporating both its theoretical underpinnings and practical applications, is necessary for assessing its status and impact across disciplines (Luhmann and Burghardt). Vanathi explores the integration of Digital Humanities into literary studies, emphasizing the potential of computational tools like text mining, machine learning, and natural language processing to analyze language patterns and uncover new dimensions within literary works. The researcher also discusses the importance of ethical considerations in Digital Humanities, particularly the responsible use of technology in creating, sharing, and consuming content (Vanathi). Cigliano, Fallucchi, and Gerardi explore the impact of digital analysis and large language models in Digital Humanities, noting the potential for these tools to enhance research in various disciplines, including literature, history, and philosophy. They argue that the integration of these tools into the humanities offers opportunities for more comprehensive and scalable research methodologies (Cigliano et al.). Upadhyay and Upadhyay outline the future directions for Digital Computational Humanities, providing a roadmap for its establishment and sustainability in data-driven organizations. They highlight the need for further development in tools and infrastructures that can support the growing demands of Digital Humanities in academic and research settings (Upadhyay and Upadhyay). Charitopoulos et al.

focus on the application of text mining in educational contexts, specifically the analysis of free-text student responses in e-assessments. Their work demonstrates how topic modeling through algorithms like Latent Dirichlet Allocation can be employed to extract meaningful topics from student essays, offering practical recommendations for course restructuring and enhancing the effectiveness of e-assessments (Charitopoulos et al.). Thus Digital Humanities has created new ways to study literature, especially when it comes to exploring how language works in texts. In Arundhati Roy's novels, the use of vocabulary carries deep emotional, cultural, and political meanings. Tools like Voyant and Online-Utility.org make it easier to look at these patterns closely, even for those without technical backgrounds. By comparing these tools, this study shows how digital methods can help readers see repeated words, themes, and shifts in language that might be missed in traditional reading. This approach adds a fresh layer to literary analysis and opens up more space for understanding how writers like Roy use language to shape their stories.

3. Methodology

This research applies two Digital Humanities tools, Voyant Tools and Online-Utility.org Text Analyser, to perform vocabulary analysis. This research adopts a comparative, tool-based digital approach to analyse vocabulary patterns Both quantitative and qualitative insights were drawn using these tools, supporting a vocabulary-centred investigation of literary style and complexity. The novels were downloaded from digital sources. This study adopts a comparative digital text analysis approach rooted in corpus linguistics to examine vocabulary patterns using two online tools: Voyant Tools and Online-Utility.org (Text Analyzer). A purposive selection of text samples was made from Arundhati Roy's The God of Small Things and The Ministry of Utmost Happiness, focusing on chapters that are thematically and stylistically representative. No pre-processing scripts or file uploads were used, in keeping with the study's intention to mirror the experience of humanities scholars who rely on graphical interfaces rather than coding environments. The texts were handled as plain prose to ensure consistency and to facilitate accurate term frequency and readability evaluations. The text was uploaded into both tools separately for analysis. Voyant Tools, a web-based suite offering advanced corpus functionalities, was used to extract data such as total word count (tokens), unique word count (types), Type-Token Ratio (TTR), frequency lists, word clouds, collocational patterns, and keyword-in-context (KWIC) results. Online-Utility.org, though more limited, provided basic lexical metrics such as total and unique word count, frequency

tables, and readability indices like Flesch-Kincaid and Gunning Fog scores. Quantitative analysis was conducted by comparing metrics such as TTR (Types/Tokens), lexical density (Content Words/Total Words), and the frequency of key terms. Qualitative evaluation considered usability, visualization capability, interactivity, and the depth of linguistic features supported. A comparative matrix was constructed to systematically contrast the strengths and limitations of each tool. Through this methodological framework, the study seeks to assess the analytical depth, technical scope, and applicability of each tool in literary corpus analysis.

3.1 Voyant Tools Working Mechanism

Voyant Tools is built using Java for backend processing and JavaScript for frontend interactivity. After uploading the data as PDF files, Voyant performed tokenization, where the content was segmented into individual units and conducted term frequency analysis to identify the most commonly used vocabulary. In addition to numerical results, Voyant generated several interactive visualizations. These included the Cirrus (word cloud), which visually represented word frequency; Trends, a line graph tracking word occurrence across the narrative; and the Contextual Reader, which displayed selected terms within their original textual context. The tool's visual outputs are powered by D3.js, a JavaScript library that enables dynamic and responsive data displays. These features collectively supported a layered understanding of vocabulary distribution and thematic prominence. Voyant Tools operates as a modular, browser-based text analysis environment that processes input through tokenisation and frequency calculations to reveal linguistic and thematic patterns. Its interactive interface enables simultaneous exploration of vocabulary through tools like Cirrus, Summary, Termberry, Trends, Terms, and Collocates. These components collectively offer layered insights into lexical frequency, distribution, and contextual relationships, facilitating both macro and micro-level literary analysis without requiring programming knowledge.

Voyant's Summary tool provides the first quantitative picture of the corpus by presenting principal statistics, such as overall word count, vocabulary density, average word length, and the most frequent and characteristic terms. It acts as a diagnostic dashboard that allows researchers to gain a quick sense of a text's lexical composition. In literary analysis, the tool is particularly valuable in recognising linguistic density, expression simplicity or complexity and initial thematic tendencies. Although it is not interpretative, it facilitates the construction of the text's stylistic orientations and sets the stage for further textual inquiry.

Voyant's Summary of The God of Small Things Vocabulary Density: 0.111

Readability Index: 8.361

Average Words Per Sentence: 11.1

Most frequent words in the corpus: said (678); like (482); rahel (434); ammu (412); estha (394); kochamma (307); chacko (299); small (279); baby (274); things (256); man (211); god (193); little (185); sophie (176); comrade (170); mammachi (164); mol (161); roy (158); arundhati (158); old (149); eyes (144); velutha (140); looked (138); didn't (131); house (130); time (121); away (120); way (119); red (119); margaret (117); just (112); ayemenem (111); pillai (109); white (106); children (106); river (105); long (105); people (100); came (100); come (98); knew (97); hair (96); took (95); went (94); room (94); left (93); thought (92); love (91); water (86); say (86); couldn't (86); day (85); years (83); know (81); head (81); wasn't (80); father (80); saw (79); look (79).

Voyant's Summary of The Ministry of Utmost Happiness

Vocabulary Density: 0.157

Readability Index: 1185.941

Average Words Per Sentence: 16.6

Most frequent words in the corpus: like (397); said (362); people (251); tilo (238); anjum (226); old (211); time (181); musa (180); way (174); knew (160); day (160); man (154); just (150); know (149); little (148); years (140); night (138); didn't (135); came (133); saddam (128); went (118); new (117); city (115); naga (114); took (110); left (108); eyes (108); told (107); it's (106); singh (104); say (104); room (102); mother (102); amrik (102); called (101); men (100); house (100); police (99); don't (98); long (97); kashmir (97); away (97); looked (94); come (93); wasn't (92); young (91); life (91); tell (90); small (90); miss (89); good (89); turned (88); army (87); body (86); head (83); asked (83); killed (81); began (81); home (80).

3.2 Cirrus

The Cirrus Component creates a word cloud in which the size of every word is a function of its frequency in the corpus. Such a visual representation facilitates an intuitive understanding of prevalent lexical items, which can point towards recurring motifs or thematic preoccupations. In literary analysis, this tool enables scholars to pick up on surface-level patterns that might merit further exploration. However, the tool only addresses frequency without contextual subtlety, so additional interpretation is dependent on close reading or ancillary tools.
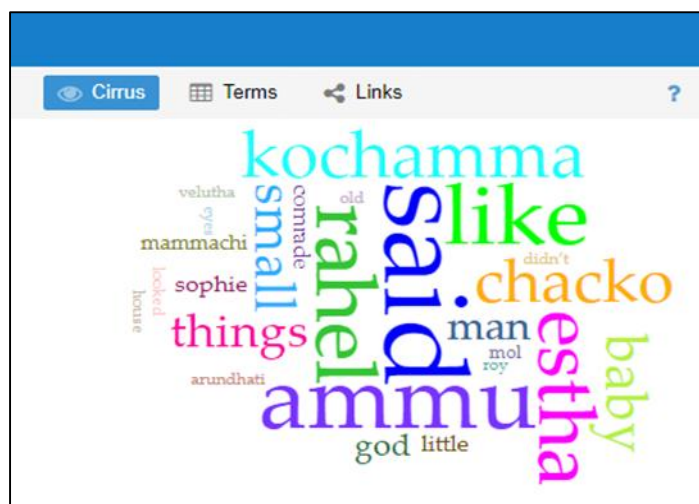
Figure 1. Word Cloud of The God of Small Things

Figure 1 explains the word cloud of The God of Small Things.
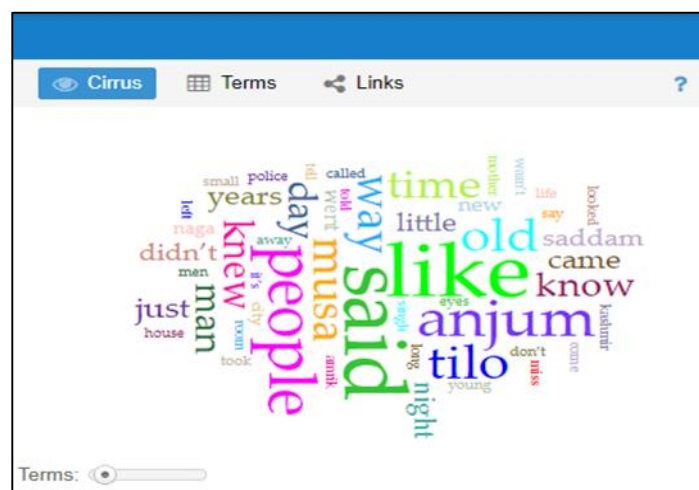
Figure 2. Word Cloud of The Ministry of Utmost Happiness

Figure 2 explains the Word Cloud of The Ministry of Utmost Happiness.

3.3 Termsberry

TermsBerry visually displays the relationship between a central word and its co-occurring terms in a network-style cluster. It not only represents frequency but also displays how words are related spatially in the corpus, which is useful for tracing the diffusion of thematic clusters or associative meanings. This graphical mapping helps find conceptual connections in literary works, facilitating thematic analysis, intertextual research, or motif tracing. Its limitation is in its focus on proximity as opposed to syntactic or semantic relationships, which demands a critical eye for interpretation.

3.4 Trends

The Trends tool graphically displays the frequency of chosen words throughout the text's timeline. It is especially useful for exploring how specific words become more or less prominent, often matching changes in story focus, thematic progression, or emotional drift. This interactive graphic works particularly well in literary analysis to monitor plot development, symbol flux, or recurring thematic issues. Although it has strengths, the Trends tool can oversimplify variation in meaning, which requires attention when read against quantitative findings.

3.5 Collocates

The Collocates feature analyzes the frequency with which particular words occur adjacent to a target keyword, revealing lexical proximity patterns. This is essential for exploring how language generates meaning in context, especially within character dialogue, emotional framing, or ideological formulations. For example, repeated co-occurrence of emotionally significant words adjacent to a character's name may indicate narrative bias or thematic focus. Although this resource provides ample information regarding contextual use, it should be interpreted carefully, as intense proximity does not necessarily mean intense semantic association.

3.6 Online-Utility.org Text Analyzer Working Mechanism

Online-Utility.org Text Analyzer is a linguistic analysis tool built with Java in HTML. This tool applies a set of predefined readability formulas, including the Flesch-Kincaid Grade Level, Gunning Fog Index, SMOG Index, and Automated Readability Index (ARI). These indices assess how difficult a text is to read based on sentence length, word length, and syllabic structure. In addition to readability, the tool provides a breakdown of word and sentence averages, lexical density, and word frequency. It also highlights "difficult" or polysyllabic words, which may pose comprehension challenges to the average reader. While it lacks graphical outputs, the tool offers a structured, easy-to-interpret set of metrics that quantify surface-level linguistic complexity. Online-Utility.org Text Analyzer employs a form-based structure that delivers lexical statistics and readability measures through automated text processing. Upon submission, the tool segments the input into analyzable units to generate data on word frequency, vocabulary density, readability, and common errors. With tools such as the

Word Frequency, Vocabulary Builder, Lexical Density, Readability Test, and The Misspellizer, it offers accessible, non-visual outputs ideal for foundational vocabulary analysis, especially for users without technical expertise.

3.7 Word Frequency List

The Word Frequency List lists each word in the text by frequency of occurrence, providing a straightforward glimpse into lexical repetition and thematic stress. This is the core of vocabulary analysis since it identifies the most prominent words utilized by the author. It allows researchers to track recurring motifs, evaluate patterns of diction, and identify often-referenced characters or ideas. Although it gives raw frequency without context, this quantitative list provides the foundation for comparative or thematic analysis between literary texts.

3.8 Lexical Density (from Readability Test)

Lexical Density is obtained from the Readability Text and reflects the proportion of content words (nouns, verbs, adjectives, adverbs) to the overall number of words in a text. Lexical density is an indicator of information richness and style intensity. Increased lexical density tends to reflect conversational or expository styles. In literary research, this method is used to distinguish narrative voices, measure stylistic variation within chapters, and rate the expression complexity.

4. Finding and Discussion

The data given by voyant tools and online.Utility.org where decoded and the result was analysed to make the comparative study. In orded to analyse the vocabulary structure of the novel three elements were considered as the basis for comparison, frequently repeated words, collocation, and lexical density. Figure 3 shows the top five words given by voyant tools termburry setting out of these five words two words the name of major characters in the novel and the other theme denote the narrative style of the novel words like said like denotes that the narrative is a third person narrative by analysing the top fifteen words, as shown in figure 4, it can be seen that the words given for God of Small Things are transliterary words and the words given for The Ministry of Utmost Hapiness can be decoded as pertaining to post-colonial literature. When the same text is uploaded in utility online-utility.org, it does not give proper

result as there is no tokenization feature available, because of which error is detected in the data generated.
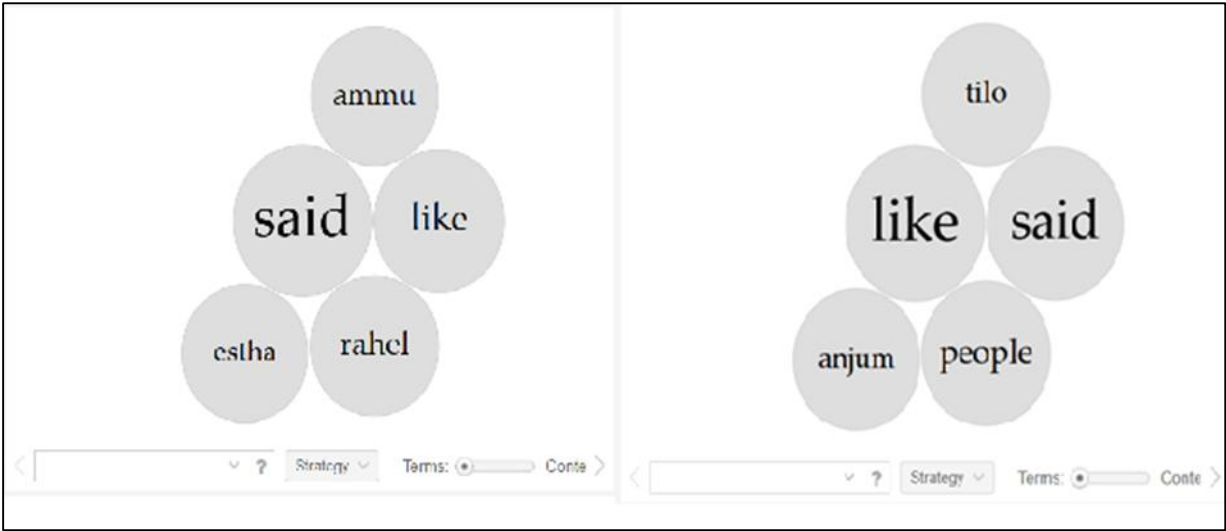

Figure 3. Top Five Words Given by TermsBerry

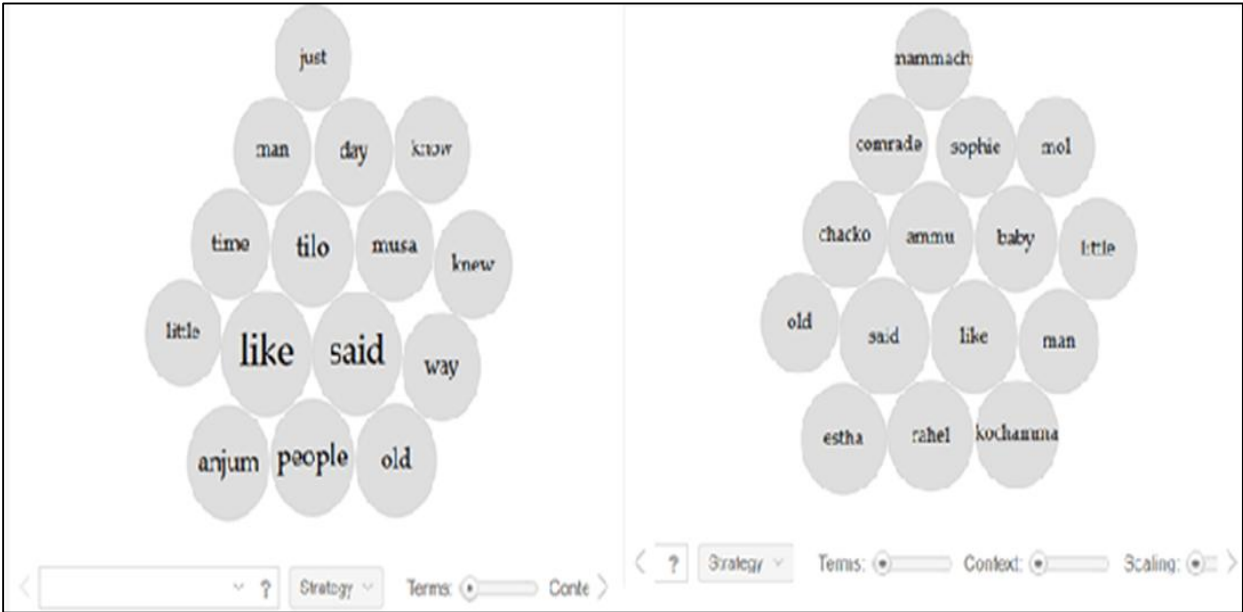Figure 3 explains the Top Five Words given by TermsBerry.

Figure 4. Top Fifteen Words Given by TermsBerry

Figure 4 explains the Fifteen Words given by TermsBerry. Another important feature in analyzing the lexical built of the novel is collocation by analyzing the relation between two words and how they appear in the novel. The researcher can understand the tone of the novel.

Figure 5 represents the tabular format of "The God of Small Things".
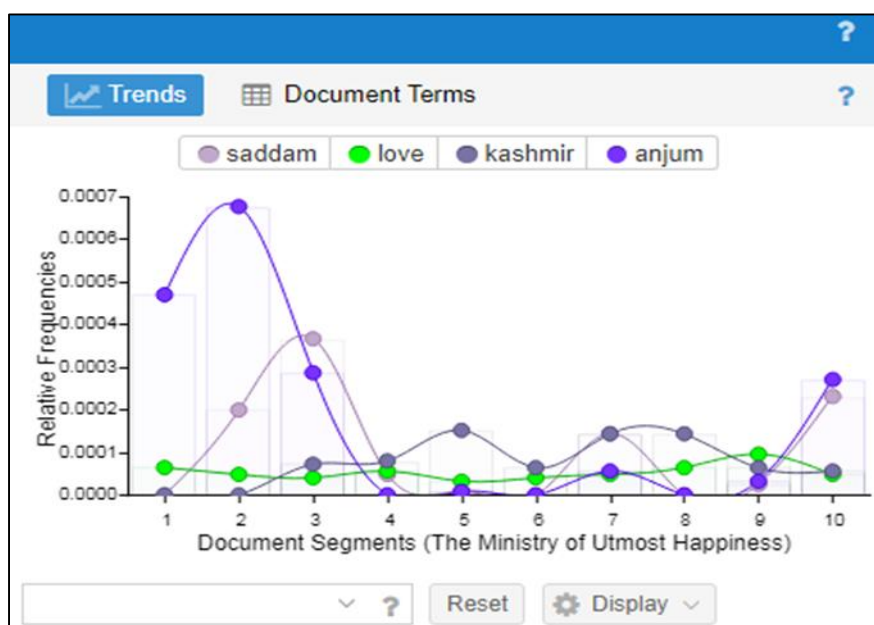

Figure 5. The God of Small Things

Figure 6. Top Words Given by Online.Utility.org for the Novel The Ministery of Utmost Happiness

Figure 6 represents the Top words given by Online.Utility.org for the novel The Ministery of Utmost Happiness.

The collocation feature provided by voyant is visually clear and statistically accurate (Figure 7 and 8) eventhough the data given by Online Utility.org is statistically accurate it requires further data visualization processing in order to analyze the text and arrive at the required results. One minor drawback in voyant tool is it does not calculate lexical density instead gives vocabulary density this is a minor defect which can be corrected by using vocabulary density. (Figure 11 and 12).
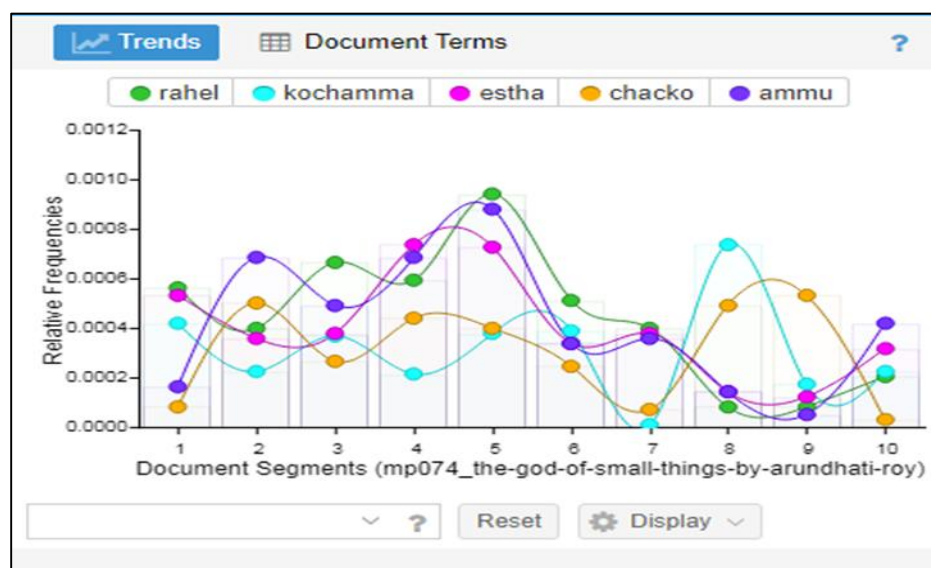
Figure 7. Trends of The God of Small things

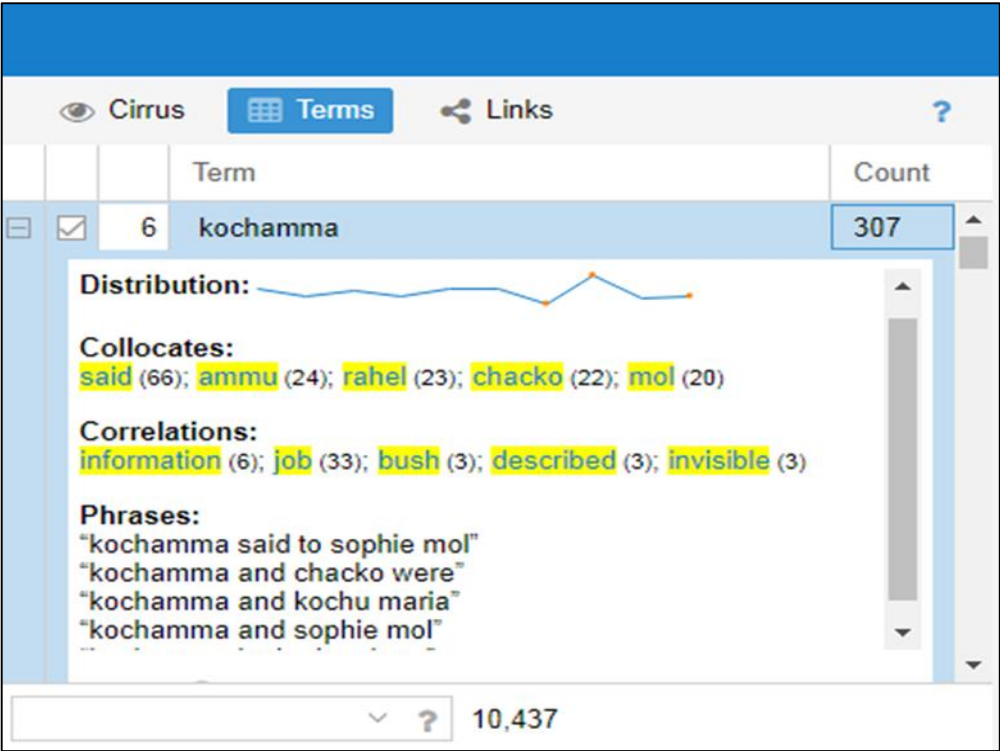**Figure 8.** Trends of The Ministry of Utmost Happiness

Figure 9. Collocated Terms of The God of Small Things

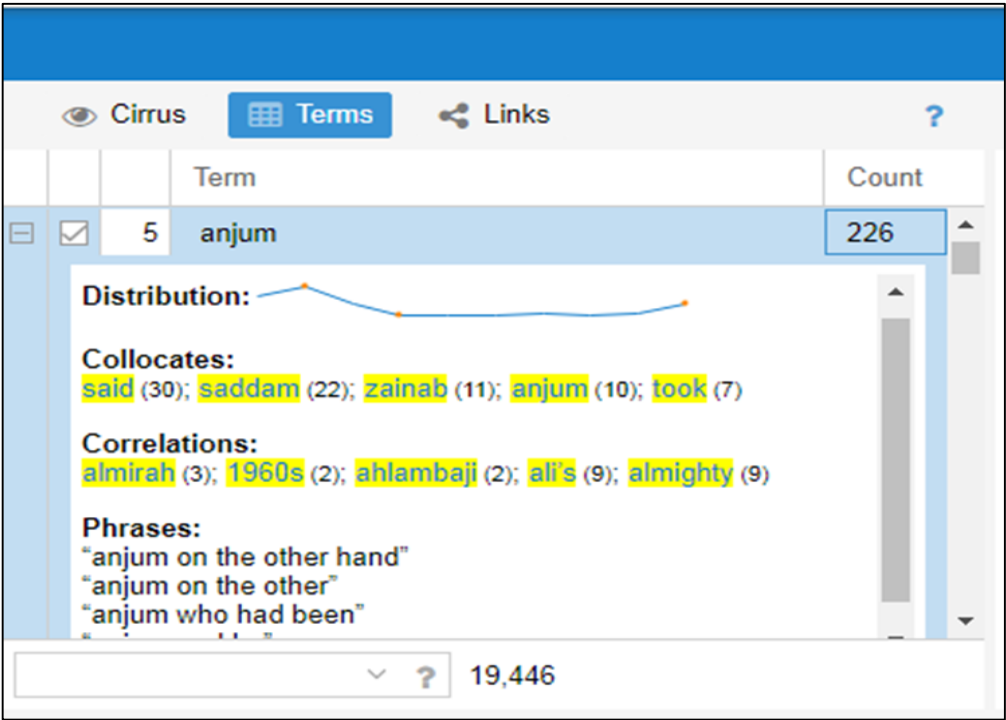Figure 9 explains Collocated Terms of The God of Small Things.

Figure 10. Collocated Terms of The Ministry of Utmost Happiness

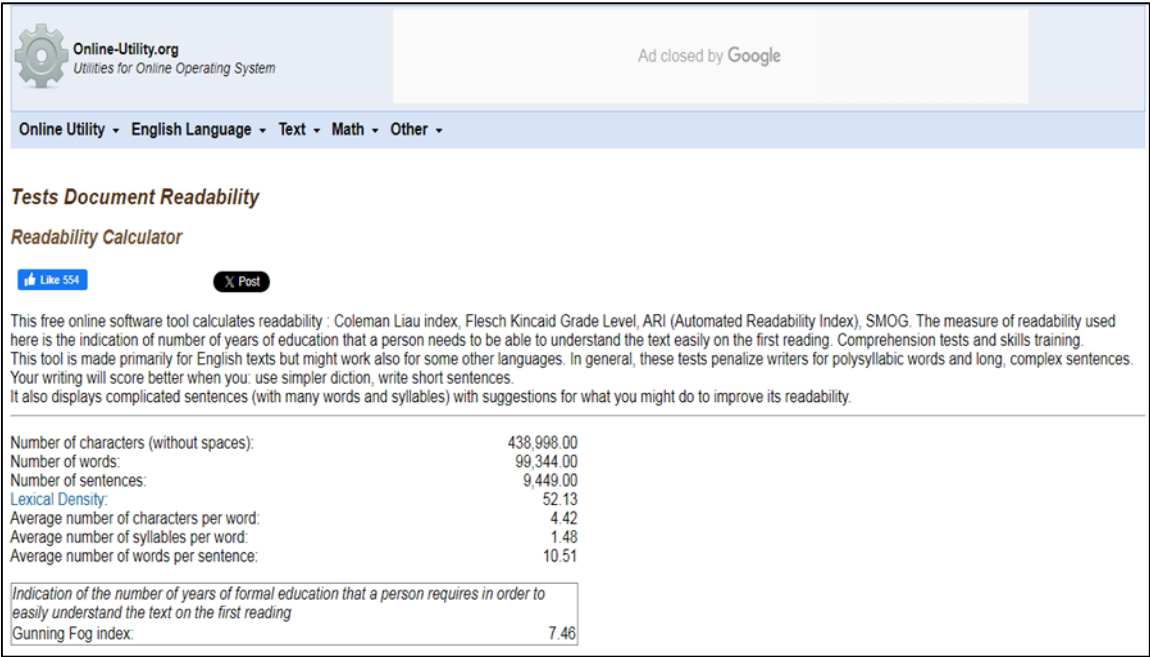Figure 10 explains Collocated Terms of The Ministry of Utmost Happiness.

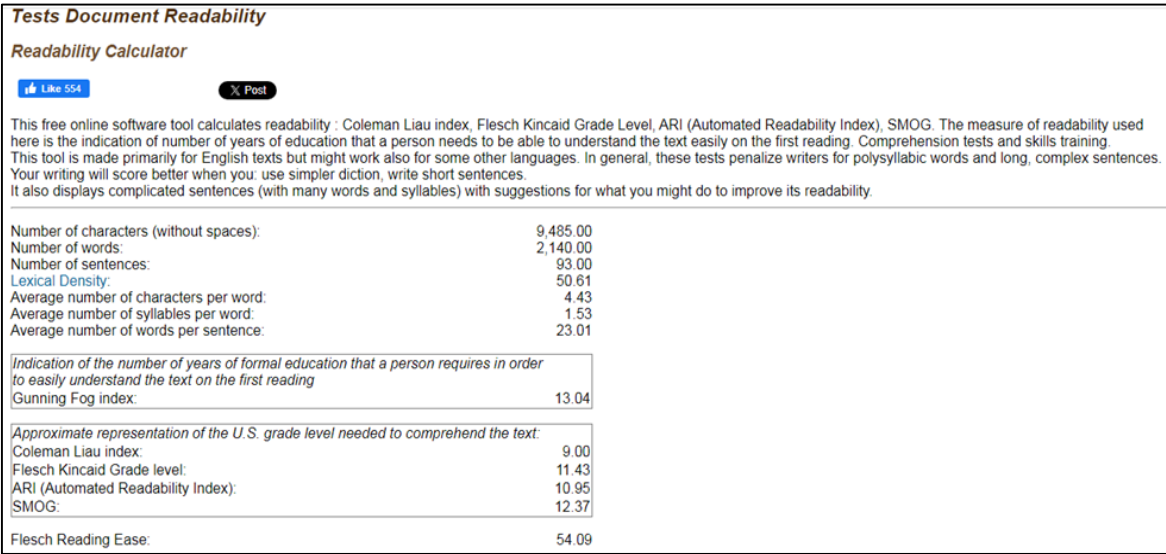**Figure 11.** Readability Test of The God of Small Things

Figure 12. Readability Test of The Ministry of Utmost Happiness

The comparative table underscores the superior analytical potential of Voyant Tools over Online-Utility.org Text Analyzer in terms of both functionality and research utility. Voyant Tools offers a visually engaging, multi-paneled interface featuring tools such as Cirrus (word cloud), Terms (word frequency), Trends (distribution across text segments), and Contexts (co-textual examination), making it a powerful tool for in-depth, interpretive analysis of large corpora. These features enable researchers to uncover thematic patterns and linguistic trends that go far beyond surface-level vocabulary statistics. Although its interface may initially seem complex to new users, with minimal acclimatization, Voyant's robust capabilities far

outweigh its learning curve. In contrast, Online-Utility.org prioritizes user-friendliness through a minimalist, text-based setup that provides basic statistics like word count, lexical density, and average sentence length. While this is useful for quick, general assessments, its limitations are significant in a research context. The tool lacks filtering capabilities for function words (articles, prepositions, auxiliary verbs), which compromises the accuracy of lexical profiling. Furthermore, the absence of data export options and visual analytics restricts the tool's applicability for comparative or reproducible research. Online-Utility.org, though efficient for rapid, one-off checks, does not support the interpretive depth or scalability required for scholarly work. Overall, Voyant's integration of visualization, and co-textual patterning makes it a more versatile and research-oriented tool than Online-Utility.org, whose simplicity is better suited to non-specialist, surface-level tasks.

Table 1. Comparison of Features: Voyant Tools vs. Online-Utility.org Text Analyzer

Feature	Voyant Tools	Online-Utility.org Text Analyzer
Interface	Visual, dynamic, dashboard-style with multiple windows	Simple, text-heavy, minimalist layout with no distractions
Ease of Use	Requires some familiarity with tools and layout	Straightforward, beginner-friendly
Word Frequency Analysis	Yes (via "Terms" tool), with filtering options	Yes, clearly listed with word counts
Word Cloud	Yes (via "Cirrus"), visually appealing	No
Trends Over Text	Yes (via "Trends"), shows frequency across text segments	No
Contextual Usage	Yes (via "Contexts"), helps in understanding word relationships	No
Lexical density	No direct output, must be inferred	Yes, the percentage provided
Average word Sentence Length	No	Yes, calculates and displays
Export/Copy Result	Can export reports in JSON/CSV format	No Export option

Best Suited for	Visual exploration, thematic analysis, timeline-based trends	Quick numeric profiling, direct comparison of vocabulary

Table 1 showing Functional Comparison of Voyant Tools and Online-Utility.org Text Analyzer for Vocabulary Analysis.

5. Conclusion

Words are the basic units of a novel. Close analysis of novels using vocabulary- analysis software like Voyant Tools and Online-Utility.org Text Analyzer highlights the importance of word structure. The comparative evaluation of Voyant Tools and Online-Utility.org Text Analyzer underscores the contrast in their functional design and practical relevance for vocabulary analysis in literary studies. After analysing the findings of the study, it was found that Voyant is more efficient in comparison to ToolsOnline-Utility.org. This approach proves especially valuable for literary scholars who require accessible, focused, and interpretable data. Voyant Tools offers a visually enriched interface with features such as Cirrus, Terms, Trends, and Collocates, making it effective for examining large corpora through thematic mapping and word distribution. Its multi-layered dashboard, ability to interpret complexity, and overlapping windows provide a better understanding of the layered textual structure. Tools like Collocates and Contexts are academically rich, and easy to navigate. Online-Utility.org Text Analyzer is a simple tool that gives raw data without any visual representation. This makes the analysis of data a mundane task. Moreover, since the text is not tokenizes, the data given is not accurate. A shared drawback of both platforms is the absence of an export function. This limitation affects data preservation and restricts researchers from saving or sharing results for further analysis or citation, which is a practical concern in academic workflows that prioritize documentation and replicability. Even with this limitation, the overall balance of simplicity, Visual output, comprehensive lexical output, and research-oriented functionality places. Voyant Tools ahead of Online-Utility.org Text Analyzer as the most suitable and effective resource for vocabulary analysis in literary research.

References

[1] Rieder, John, and Theo Röhle. "Digital Humanities Manifesto." Digital Humanities: A Primer for Students and Scholars, 2009.

[2] Moretti, Franco. Distant Reading. Verso, 2005.

[3] Liu, Alan. "Where Is Cultural Criticism in the Digital Humanities?" Debates in the Digital Humanities, 2012.

[4] Aathira, S., et al. "Digital Humanities in Language and Literature: Unveiling New Worlds of Inquiry." Educational Administration: Theory and Practice, vol. 30, no. 11, 2024, pp. 335–339.

[5] Knight, David, et al. "FreeTxt: A Corpus-Based Bilingual Free-Text Survey and Questionnaire Data Analysis Toolkit." Applied Corpus Linguistics, vol. 4, 2024, p. 100103.

[6] The Challenge in Analyzing Vocabulary." Digital Scholarship in the Humanities, vol. 36, suppl. 2, 2021.

[7] Luhmann, Jannis, and Mario Burghardt. "Digital Humanities—A Discipline in Its Own Right?" Journal of the Association for Information Science and Technology, vol. 73, 2022, pp. 148–171.

[8] Vanathi, A. "Exploring Literary Studies through Digital Humanities." Shanlax International Journal of English, vol. 12, no. S1, 2023, pp. 203–207.

[9] Cigliano, Andrea, Fabio Fallucchi, and Massimiliano Gerardi. "The Impact of Digital Analysis and Large Language Models in Digital Humanities." ICYRIME 2024, Catania, 2024.

[10] Upadhyay, Sanjay, and Nidhi Upadhyay. "Future Directions and a Roadmap in Digital Computational Humanities." ITQM 2017, Elsevier, 2017.

[11] Charitopoulos, Andreas, et al. "Text Mining Technologies Applied to Free-Text Answers of Students in E-Assessment." Journal of Educational Technology & Society, 2025. https://doi.org/10.1007/s10791-024-09496-9.

[12] Roy, Arundhati. The God of Small Things. IndiaInk, 1997.

[13] Roy, Arundhati. The Ministry of Utmost Happiness. Penguin Random House India, 2017.

[14] Sinclair, Stéfan, and Geoffrey Rockwell. "Voyant Tools." Voyant Tools, 2016, https://voyant-tools.org/.

[15] Online-Utility.org Text Analyzer." Online-Utility.org, https://www.online-utility.org/english/readability_test_and_improve.jsp