A Corpus Stylistic Analysis of Some Lexical and Semantic Devices in Young Adult Dystopian Fiction

Azza Abdel Fattah Abdeen
Faculty of Women
Ain Shams University

Abstract:

This research examines some lexical and semantic devices such as collocation, semantic preference, semantic prosody and lexical clusters in a compiled corpus of four trilogies of young adult dystopian fictions that cover the period from 2008 to 2013. The AntConc software is used for the Word List, Collocation, N-Grams and Concordance tools. The employed linguistic framework is eclectic because it draws on techniques and models from Mahlberg (2007), Leech and Short (2007), and Sinclair (2007). The main objective of this study is to illustrate how the techniques of corpus linguistics validate literary analysis. Findings of the study revealed that corpus linguistics aids the exploration of textual and thematic features that may transcend traditional stylistic analysis. Findings also showed that frequent occurrence of nouns referring to different parts of human body is associated with the important role played by body parts in dystopian fiction to identify the characters' physical and habitual traits. Furthermore, the frequent occurrence of adjectives as collocates of nouns referring to body parts reveals that dystopian fiction is descriptive and visual in nature.

Key Words: corpus stylistics, collocation, semantic prosody, lexical clusters, body parts clusters, dystopian fiction.

A Corpus Stylistic Analysis of Some Lexical and Semantic Devices in Young Adult Dystopian Fiction

Azza Abdel Fattah Abdeen
Faculty of Women
Ain Shams University

1. Introduction:

Thanks to recent advances in computational linguistics, the study of significant linguistic notions such as collocation and lexical clusters has become accessible with regard to large corpora. Computational linguistics has equipped researchers with tools that enable the investigation of large compiled data of literary texts yielding reliable results upon which qualitative interpretations are built. These notions of collocation, semantic preference, semantic prosody and lexical chains have attracted much research in different areas; however, little research has been done to apply these linguistic notions to literature. Thanks to the help of corpus linguistics tools, the retrieval and investigation of some of these elusive semantic notions across corpora have become manageable. This study adopts a corpus-driven approach to literary texts with regard to register together with a reader-based approach. To this end, the study employs an eclectic framework based on Leech and Shorts' checklist (2007) for the analysis of the different types of adjectives, Mahlberg's clusters (2007) for the analysis of lexical bundles, and Sinclair's model (2004) for the analysis of collocation, semantic prosody and semantic preference. These lexical cohesive devices facilitate the process of narration in fiction making it sequential and understandable by the readers. Cohesion, in general with its two parts: grammatical and lexical, gives a text its texture by uniting its parts together for the text recipient to make sense out of it. This study investigates the role of the lexical cohesion devices:

collocation, colligation in explicating the stylistic characteristics of a specific register. Semantic prosody, semantic preference and lexical clusters are investigated within the lens of collocation. Intricate textual semantic relations are made eligible using corpus software.

This study also investigates the textual features that play a major role in creating the popularity and sense of fantasy of dystopian fiction, marking it as a register. The study relies on corpus linguistics for quantative analysis of the selected fictions, yet "necessitate(s) insights from other areas of stylistics for a full elucidation of their stylistic effects" (McIntyre, 2015, p. 60). The free AntConc software is used for Word list, Collocation and N-Grams and Concordance tools to show how corpus tools can help provide the lexical categories that aid a further qualitative analysis. The quantitative approach and the qualitative analysis are meant to throw light on the crucial role these devices play in text unity and interpretation. Analyzing a text in terms of its lexicon may disclose recurrent themes that might not be easily otherwise revealed. "Lexical feature is concerned with the choice of specific lexical items and their distribution and relation among themselves, and the way they generate meaning in the text" (Faiz, Azher, Musarrat, Asghar, & Jabeen, 2018, p. 330). Investigating the lexical construction of literary texts can unleash significant cohesive devices that might reveal dominating themes.

1. 1. Young Adults Fiction (YAF):

The emergence of YAF as a recent genre has attracted research work. However, a few linguistic studies have searched YAF as a genre so far. Even fewer studies have applied corpus linguistics as a methodology to fiction subjecting it to quantitative and qualitative analysis. YAF as a recent genre started growing public interest with the famous *Harry Potter* series (1997) and continued with The Hunger Games series by Suzanne Collins (2008). The 2000s has

witnessed a more flourishing period of this genre. It refers to that kind of literature that links children literature to adult literature. Being a recent genre that relates to a short transitional stage in one's life, it has not received much academic interest. There are hardly any interdisciplinary studies on the language of young adult literature. Nelson (2016) states that "because the emergence of young adult literature as a recognized genre is a fairly recent phenomenon, relatively few studies have specifically targeted YAF. Nelson (2016) states that "while there is certainly room for more YA research across disciplines, there has been more research on YA produced by literature and education disciplines than linguistics." (p. 5).

 YAF has been labelled as Fantasy fiction and juvenile fiction. The reader of the YAF is an adolescent of 14-18 years old, a period where a child starts developing symptoms of physical and mental maturity. Therefore, the main characters of YAF should also fall within this age. They have characteristics that appeal and meet the needs of a young adult. Love, friendship, family relationships, death, and self- exploration are major themes in YAF.

1.1.1. Dystopian Fiction

Dystopian fiction is a common and popular register of young adult literature. Register is used to refer to language variation "of a nondialectal type; e.g. differences between polite and familiar language; spoken and written language; scientific, religious, legal language, etc." (Leech and Short, 2007, p. 65). According to Oxford dictionary, dystopia is a made-up word of the prefix "dis" which means bad and "utopia" which means "an imagined place or state of things in which everything is perfect" (https://en.oxforddictionaries.com, n.d.). Dystopia is the opposite of the utopian world as it represents the imperfect and corrupted world. It means "an imagined state or society in which there is great

suffering or injustice, typically one that is totalitarian or post-apocalyptic" (1, n.d.). Utopia, in contrast, represents the perfect, idealistic dream world. Dystopia and utopia are the two sides of the same coin; an imaginary world of either a bright side (utopia) or a dark side (dystopia). A dystopian society stands in striking contrast to the idealistic, orderly life represented by the utopian society. It represents a society where its inhabitants rebel against regulations, reject orders, and revolt against external power. They are fostered by the urge of changing their existing world into a better one. They dream of a world of their own creation. The main hero in dystopian fiction is a teenager who is willing to strive a journey into the unknown.

> They cross the threshold into a world of darkness and danger, of allies and enemies, and begin a journey towards their own destiny that will change their world. They will be tested, often to the very edge of death. The stakes are high. The adults are the oppressors. The children are the liberators. It's heady stuff, far removed from the routine of everyday life. (Young, 2011)

Dystopian fiction belongs to science fiction genre because "many of them use other elements of science fiction like time travel, space flight, amazing and advanced technologies, etc." (Utopia and Dystopia, 2018). Many Dystopian novels are written as trilogies with a different title assigned to each book. Dror (2014) argues that "Dystopian fiction is a productive place to address cultural anxieties and threats as well as to contemplate the ideal or the utopia" (p. 6). Dystopian fiction describes a fantasy world, a world different from the authentic and visible one. The word fantasy itself evokes "images of witches, fairies, dark woods, magic wands and spells, time travel, ghosts, and dragons" (Gates, Steffel, & Molson, 2003, p. 1). It "sets youth at odds with adults and empowers young people to turn against the system" (Dror, 2014, p. 8). It takes good versus evil as its dominant theme. The main protagonists in their

everlasting fight against evil continually try to defeat treacherous characters or powers. Similarly, the theme of survival is quite recurrent in dystopian fiction. The characters, whether good or bad, possess super powers that intensify the struggle between good characters and bad ones and enable them to finally win the battle and restore the utopian world. Fantasy, in this case, is quite appealing to adolescents because, at this life stage, they start growing passion for exploring worlds other than the real one. They also start developing growing interest in invisible things. This is part of their emotional and mental growth. Hence, fantasy, with its mythical world, fosters their imagination. Perhaps, they find release in escaping their realistic environment to a mythical one that may have more potentialities for them. This could be one way of seeking self fulfilment by transforming their world into a more exciting and mysterious one. Fantasy, in this sense, helps young adults mature. Dror (2014) believes that "dystopian young adult texts focus, on the whole, on the actions of humans in the developed western world. They comment on an increasing reliance on technology, genetic modification, consumerism, and the massive environmental impact this lifestyle causes." (p. 8).

1.2. Objectives of the Study:

This research is concerned with how solid semantic relationships that hold the text together can give rise to recurrent themes and stylistic features that may not be easily explored through traditional analysis.

The objectives of this study are:
1. To demonstrate how corpus linguistics validates a quantitative and qualitative analysis of literary texts.
2. To show how the lexical choices of the writers of dystopian fiction is related to common themes in the fictions.

3. To examine the relation between the semantic prosody and semantic preference of the most recurrent lexical items and their stylistic effect in communicating some personal traits of the characters.

1.3. Hypothesis and Research Questions:

This research is set to test the validity of the following hypothesis:
1. Corpus linguistics as a methodology validates quantitative and qualitative analysis of literary texts.
2. Writers of dystopian fiction choose their lexis in a way that helps evoke a particular effect in their target readers and give the fiction its solid texture.

The current research addresses the following questions:

1. How does the analysis of collocation lead to the exploration of recurrent themes in the corpus of the study?
2. How can the most frequent words retrieved from the word list tool indicate recurrent themes in dystopian fiction?
3. What are the types of adjectives that collocate with nouns in the corpus?
4. What function do body parts have in dystopian fiction?

1.4. Significance of the Study:

The significance of the study stems from the fact that it merges three distinct fields of study: literature, linguistics and corpus linguistics. The three areas collapse into a stylistic analysis using recent advances in corpus linguistics. Hence, it may arouse the interest of students of literature and linguistics to see how these areas are interrelated for a stylistic analysis of literary works based on finding evidence in the text itself.

2. Review of Linguistic Models and Techniques:

This section reviews the main linguistic techniques that form the backbone of the linguistic model to be applied to the corpus of the study. It also reviews the studies that are found to be related to the current one in an attempt to find out how similar/different this study is to/from them.

2.1. Corpus Linguistics:

Corpus linguistics has evolved from computational linguistics to test the applicability of computational software tools against naturally occurring data. It has developed as an objective technique to subject large corpora to an objective investigation. Therefore, reliability and replicability are fundamental to corpus linguistics. Its main concern is about finding systematic patterns in language and decoding the meanings attached to those patterns. According to Fischer-Starcke (2010) corpus linguistics makes it possible to (1) run a systemic lexical and grammatical analysis of a large collection of data looking for patterns, (2) interpret the meanings in those patterns (p. 1). It "provides data with which conclusions can be reached more consistently" (Sinclair, 2010, p. 15).

This study is empirical and descriptive in nature because it looks for the linguistic evidence in the text that gives rise to generalizations. Interpretation and meaning are then drawn based on these generalizations. Corpus linguistics equips the researcher with the tools that enable him/her to do a quantitative analysis of corpora with reliability and objectivity. Researchers find and trace systematic patterns in their analytical corpus using tools such as Concordance, Key Word, Word List, Collocation and N-Grams. The Concordance tool, for example "is a display format that shows a search word with a specified amount of co-text to its left and to its right" (Mahlberg, 2010, p. 292). The Key Word tool enables" the comparison of word frequencies across corpora, annotating corpora with further information to retrieve specific grammatical patterns, and applying statistical tests to assess the significance of frequency information" (Mahlberg, 2010, p. 293). Corpus linguistics has

proved to be a valid methodology for analyzing literary texts since the linguistic evidence lies in the text itself. It also helps a researcher decide on which extracts to analyze "as well as providing a means of examining whole texts as opposed to small extracts" (Mahlberg & McIntyre, 2011, p. 206).

In short, a literary text manifests itself lexically and grammatically via linguistic resources which can be traced by using a software for a statistical analysis. This has given rise to corpus stylistics.

2.1.1. Corpus Stylistics

Stylistics in its broadest sense is the field that applies theories, principles and techniques offered by different linguistic fields and sub fields to various literary and non-literary texts. It is an approach for the study of style from a linguistic perspective (Leech and Short, 2007). It focuses on looking for the linguistic evidence that gives rise to a particular interpretation. Therefore, it borrows theories and techniques from various linguistic theories and models for a rich analysis. Corpus stylistics, on the other hand, is a coined approach to describe a stylistic analysis of a text incorporating corpus tools. It is an up-graded version of traditional stylistics that subjects literary texts to objective analysis by using a corpus software. McIntyre (2015) defines corpus stylistic as "the application of theories, models and frameworks from stylistics in corpus analysis" (p.61). Balossi (2014) argues that "when corpus linguistics involves the study of literary corpora, it is frequently referred to as corpus stylistics or corpus literary stylistics to highlight that the analysis is centered on literary works" (p. 49). In other words, it looks for the linguistic evidence that gives rise to a particular meaning in a literary text. Mahlberg (2007) argues that "the application of corpus methodology to the study of literary texts can be described as 'corpus stylistics'. Corpus stylistics investigates the relationship between meaning and form. Thus, it is similar to both stylistics and corpus linguistics" (p. 4). What differentiates corpus stylistics from corpus linguistics is the fact that the former "needs to incorporate

theories, models and methods from qualitative stylistic analysis to augment computational techniques" (McIntyre, 2015, p. 60).

Research in corpus stylistics has gone through two main streams: annotation and collocation (Wynne, 2006, p. 1). This study follows the second stream which is to explore some stylistic aspects by looking for the linguistic evidence manifested by the text. Thus, corpus stylistics has recently become a trending field for stylistic analysis because it aids the systematic analysis of corpora that may take a lifetime to be analyzed manually. Mahlberg argues in favor of this point that "corpus linguistic methods that take into account statistical measures of significance may be difficult to apply to the study of individual texts, because there may simply not be enough data to make reliable claims" (Mahlberg, 2010, p 296). The major difference between traditional stylistics and corpus stylistics is that the former focuses on finding linguistic deviation from the norm that gives rise to a stylistic value; whereas, the latter "combines the study of deviations that define artistic expression and inform us of the text's aboutness or content, with the study of recurrent patterning of language that can give rise to unique stylistic findings. Frequencies of deviated words and patterning are identified using computer software" (Spencer, 2011, p. 7).

Corpus stylistics helps researchers gain insights into the text by investigating the lexical and/or grammatical choices of the author of the literary texts. McIntyre (2013) suggests that approaches to corpus stylistics may be classified as (i) corpus-assisted analysis; (ii) corpus-based analysis; and (iii) corpus-driven analysis. According to McIntyre (2013), a corpus-assisted analysis depends on large reference corpora to assist the analysis of a single text. Corpus-based analysis, on the other hand, is set to test the researcher's hypothesis about the stylistic effects of the writer's linguistic choices by uploading the target text to a software. Corpus-driven analysis, in contrast, uses the results of corpus analysis to formulate new theories and hypotheses (McIntyre, 2013). According to

McIntyre's (2013) classification, this study is a corpus-based analysis as it is set to test the stylistic effect of the writers' lexical choices. In general, corpus stylistics offers researchers with statistics and tools that make their analysis more objective and reliable. No wonder then, that corpus stylistics has achieved great leaps as a sub discipline of corpus linguistics to the extent that it may be considered as a field in itself.

Many studies have contributed to the development and elaboration of corpus studies e.g., Mahlberg (2006, 2007, 2010, 2011); Leech and Short (2007); Stubbs (2005). Most of them analyze lexicogrammatical patterns of varied literary texts using quantitative and qualitative methods. Quantitative methods, such as word frequency, provide the researcher with a "a global insight into the linguistic or stylistic features of a text and help identify those worth considering in greater depth through qualitative methods" (Balossi, 2014, p. 51). The statistical facts provided by the corpus-based analysis are not considered significant unless they are qualitatively interpreted "to address theoretical questions" (Mahlberg, et al, 2013, p. 35). Many other studies focus on textual and lexical features that characterize a particular literary text or writer (Biber, 2011, p. 15). Most studies rely for their analysis on key word analysis and collocation analysis and lexical phrases that characterize a particular literary work or author. Corpus stylistics facilitates the process of "understanding of the linguistic units in literary texts and the effects these have on the way in which readers create meanings from texts" (Mahlberg et al, 2013, p. 36).

2.2.   Collocation:

Lexical cohesion "concerns the way in which lexical items are related to each other and to other cohesive devises so textual continuity is created" (Flowerdew, J and Mahlberg, 2009, p. 1). It contributes to the cohesion of a text through the selection of vocabulary" (Al-Pachachi, and Nase, 2016, p. 36). Lexical cohesion

also contributes significantly to the interpretation process of a text. Giannossa (2012) comments that "the author of a literary text chooses his or her lexis having a specific purpose and readership in mind. Through word choice, s/he aims at evoking a particular emotion or reaction on the reader's part (Giannossa, 2012, p. 3).

The cohesive devices that achieve lexical cohesion are reiteration and collocation. Reiteration means "the repetition of a lexical item or the use of a synonym of some kind in the context of reference" (Halliday & Hasan 1976:318). Collocation as a technical term is first introduced by Firth (1957). Collocation in its simplest form refers to the words that tend to occur together. Linguistically speaking, it refers to lexical and grammatical patterns with semantic relations that determine their meanings. Biber (2011) defines collocation as a systematic association of certain words to occur together frequently (p.18). It consists of a node word and accompanying collocates. In corpus linguistics, collocation is basically described as the actual occurrences of words in texts. Sinclair (2004) originally defines it as "the cooccurrence of words with no more than four intervening words" (p. 141). In corpus linguistics, in order for a set of lexical items to be identified as collocates, they need to appear in the text with a certain frequency and the number of intervening words between them should not exceed four. For Halliday and Hasan (1976), by contrast, textual evidence is not fundamental; what matters is the meaning associations between words (Flowerdew & Mahlberg 2009: p. 112). Mahlberg (2009) believes that "meaning is distributed over several words that are chosen together. The way in which words are chosen together can be observed in collocational patterns: the co-occurrences of words in texts. Such patterns create cohesion" (p.112).

Collocation in corpus linguistics "refers to an individual lexical item that is found through statistical measure to co-occur significantly

with another lexical item" (Wijitsopon, 2013, pp. 45-46). Indeed, corpus linguistics has facilitated the analysis of collocation in large corpora. Philip (2011) states that computers contribute to collocation studies in two ways: (1) retrievability, (2) quantification. Retrievability helps extract nodes in context (p.43).

In corpus linguistics, collocational strength is measured by a number of statistical tests such as " MI (mutual information), z, t, log-likelihood, log-log, and MI3 scores" (Xiao, & Mcenery, 2006, p. 105). These statistical measures help "set the minimum co-occurrence frequency of an item to be considered as a collocate of a given node word" (Xiao, & Mcenery, 2006, p. 105). Begagić (2013) studies the semantic preference and semantic prosody of the collocation "make sense" using the statistical measurement MI in the Corpus of Contemporary American English (COCA). He analyzes manually all the forms of the collocation 'make sense' at the span of 10 words to the left and 10 words to the right (Begagić, 2013, p. 403).

   2.2.1. Colligation:

Collocation has a grammatical version called colligation. Unlike collocation, colligation, introduced by Firth (1957) and developed further by Sinclair (2004), refers to grammatical classes that show co-habitual occurrence with a given node. For example, the frequent occurrence of definite articles and adjectives in association with nouns are expressed as colligation. In other words, whereas collocation expresses a relationship of co-occurrence at a lexical level, colligation expresses a relationship of co-occurrence at a grammatical level. Colligation may demonstrate " preference for a particular verb tense, negative particles, modal verbs, participles, that- clauses, and so on" (Philip,2011, p. 51). It is worth saying that this study deals with collocation and colligation patterns that are found significant.

2.2.2. Semantic Preference:

Semantic preference is first used by Louw (1993) to refer to a group of words which share similar semantic features. According to Philip (2011) "semantic preference, the third element of the unit of meaning, is an abstraction of the semantic patterning which pulls together collocates and non-recurring near-synonyms into semantic sets" (p54). Sinclair (2004, p. 142) defines semantic preference as "the restriction of regular cooccurrence to items which share a semantic feature, for example that they are all about, say, sport or suffering". For example, Sinclair (2004) notices that, the node "*true feelings*" shows a strong association with possessive adjectives.

2.2.3. Semantic Prosody:

Semantic prosody is another related concept that is studied through the lens of collocation. Biber (2011) defines semantic prosody as "an underlying evaluative meaning, which is usually categorized simply as whether the object or event is considered to be good or bad" (p. 18). Semantic prosody is evaluative "and on the pragmatic side of the semantics/pragmatics continuum" (Sinclair, 2004, p. 34). Sinclair confirms that " semantic prosody has a leading role to play in the integration of an item with its surroundings" (Sinclair, 2004, p. 34). In addition to that, semantic prosody colors the lexical item by assigning to it either a positive or a negative connotation. Semantic prosody is effective in directing the tone of a text (Mahlberg, 2006, p. 374). Biber (2011) claims that "the notion of semantic prosody turns out to be very useful for explaining the stylistic effect of particular expressions in a literary work." (Biber, 2011, p. 18). Semantic prosodies of most lexical items, according to Louw (1993), tend towards negativity "with relatively few of them bearing an affectively positive meaning. However, a speaker/writer can also violate a semantic prosody condition to achieve some effect in the hearer—for example irony, insincerity, or humour can be explained by identifying violations of semantic prosody" (Louw

1993: 173). A well-known example of negative prosody is the verb 'cause' which is frequently used in negative contexts such as cause death, accident, illness. It is worth stating that semantic prosody and semantic preference are investigated in this study through the lens of collocation

2.3. Lexical Clusters

Lexical cluster is another linguistic feature that computational linguistics has made accessible to search. A cluster refers to a group of repeated words in a sequence. For a set of lexical items to be called a cluster, they should occur at least ten times together in a corpus. They are also referred to as lexical bundles and lexical chains. All these terms are used alternatively in the study. N-Grams is the AntConc tool for extracting recurrent clusters automatically. The function of N-Grams tool is to find recurrent sequence of specific items that occur together. The study of lexical bundles or clusters focuses "on general patterns that hold across a number of texts in a register or subcomponent of a corpus" (Mahlberg, 2007, p. 5). The repetition of a sequence of words in fiction has " identifiable discourse functions in texts" (Conrad, & Biber, 2004, p. 6). lexical clusters are a chain of a lexical and grammatical set with a certain semantic prosody (Mahlberg, 2007). Clusters combine lexis and grammar, hence, they have lexical and textual function in addition to their important role in creating cohesion and textual meaning (Wijitsopon, 2013, pp. 45-46).

For the purpose of this study clusters of three words are retrieved from the corpus using the N-Grams tool of AntConc. They are investigated through the lens of collocation for thematic purposes. For example, the clusters *my hand on* and eyes on me (see The Analysis section for further discussion) are a sequence of three words that occur more than ten times in the corpus of the study, hence worth investigating for their textual meaning and function.

3.   Review of Related Studies on Corpus Stylistics:

There is a large number of studies applying corpus linguistics to literary texts (Mahlberg 2006, 2007, 2010, Mahlberg and McIntyre 2011, Mahlberg et al, 2013, O'Halloran, 2007, Sinclair, 2010, Stubbs, 2005). Most of them analyze the lexicogrammatical patterns in one or two literary texts. The reason for doing this is sometimes twofold: either to trace some stylistic features of the authors of these texts and/or to trace some features characterizing a particular register. For example, O'Halloran (2007) applies a corpus approach to James Joyce's novel "Eveline" using the Key Word List and concordance lines "in order to reveal some of these subconscious intimations whilst reducing as much as possible arbitrariness and circularity in analysis and interpretation" (p. 227).   O'Halloran (2007) compares the novel to the BNC-Baby (British National Corpus) reference corpus to extract key words.

Leech and Short (2007) introduce a checklist of linguistic and stylistic categories which include lexical categories, grammatical categories, figures of speech, and cohesion and context. The lexical categories for nouns as introduced by Leech and Short (2007, p. 61), for example, enable the researcher to see whether the nouns are abstract or concrete, the kinds of abstract nouns, the use of proper names and collective nouns. The lexical categories for adjectives according to Leech and Short are based on whether the adjectives are frequent and whether they denote physical, psychological, visual, auditory, color, referential, emotive or evaluative attributes. Adjectives are also categorized according to whether they are adjectives restrictive or nonrestrictive, gradable or non-gradable or attributive or predicative (Leech and Short, 2007, p. 61). The current study adopts Leech and Shorts' classification (2007) of adjectives in the analysis of nouns and adjectives by categorizing nouns according to their semantic preference and categorizing adjectives according to the function they fulfil in the analyzed collocation patterns.

Stubbs (2005) discusses some linguistic features in Joseph Conrad's Heart of Darkness such as "patterns of phraseology in the text itself, and their relation to patterns in the language as sampled in large computer-readable corpora" (p. 5). Stubbs (2005) searches for recurrent phrases using the N-Gram tool as a way of revealing current themes in the novel.

Mahlberg (2010) comprises a corpus of Dickens' novels and a large reference corpus of all 19th C fictions. Mahlberg (2010) focuses on how corpus linguistics as a methodology validates quantitative and qualitative analysis of literary texts. For instance, Mahlberg and McIntyre (2011) use corpus linguistics for semantic analysis of the key semantic domains of the key words in *Fleming's Casino Royale*. They classify the key words into two groups characterized as text-centered and as reader-centered. The findings of their study prove that "corpus methods can aid the identification of elements of a text worthy of further qualitative analysis" (Mahlberg. and McIntyre, 2011, p. 225)

Mahlberg (2007) analyzes the repetition of five-word clusters in Dickens' novels and identifies five functional groups Labels for speech clusters: As If clusters, Body Part clusters and Time and Place clusters. Mahlberg (2007) analyzes clusters as "pointers to textual function" (Mahlberg, 2007, p. 5). She finds out that the lexical chain "his hands in his pockets" and "as if he would have" are more frequent in Dickens' novels than in the reference corpus of a 19th C. fiction. She also reaches a signifincant conclusion that the most frequent clusters are the short ones. Longer clusters of more than five-word clusters tend to be less frequent in the Dicknes corpus.

This research adopts Mahlberg's identification (2007) of clusters as recurrent sequences of words that indicate local textual functions. It also follows Mahlberg's classification (2007) of lexical clusters into functional groups and labelling them.

Mahlberg et al (2013) study clusters of body parts in Dickens' novels under the concept of suspensions. Suspensions means that

body language and speech happen simultaneously (Mahlberg, M, Smith, C and Preston, S, 2013, p. 43). According to the writers, "Suspensions often interrupt the fictional spoken language with information on the context in which the speech occurs and the lexico-grammatical patterns of suspensions support the presentation of information as contextual" (Mahlberg et al 2013, p. 43). Body clusters are used according to Mahlberg et al (2013) "to fulfil local textual functions in the creation of fictional characters" (p. 41). They reflect human features that aid characterization. They also can make the situation more authentic by making it more realistic and vital (Mahlberg et al, 2013, p. 41). They also help the reader to visualize the character and situation. " The unobtrusive character of such Body Part clusters is also reflected by the fact that they seem to have common collocations or collocating clusters" (Mahlberg M. , 2007, p. 23). Mahlberg's (2007) analysis of body part clusters in Dickens' novels is based on Korte's (1997) model on body language. Mahlberg et al (2016) developed the web application CLiC to support the identification of textual patterns in fiction from a cognitive perspective. Mahlberg et al (2016) argue that "corpus methods broaden the view from bottom-up cues in an individual text to a more general account of fictional speech patterns across texts that affect the top-down processes that are relevant to mind-modelling.

This research fits in with Mahlberg's use of lexical clusters for a contextual function. It also adopts her approach in labeling these phrasal clusters and grouping them. However, the current research adopts a different corpus tools other than extensively used by Mahlberg. It does not depend on the key word list tool as compared against a larger corpus. It extends Mahlberg work which mostly focuses on extracting the most frequent key words in the texts under analysis by focusing on the different categories of collocation for a thorough corpus stylistic analysis. Moreover, the current study does

not apply the cognitive stylistics approach to the analysis of the corpus because this is beyond the scope of the study.

Wijitsopon (2013) applies a corpus-driven approach to Jane Austen's six major novels to see how well corpus approach is applicable to literary texts. He investigates the lexical patterns and their textual function in six novels by Jane Austin. He uses the key word corpus tool in comparison to a reference corpus and uses three descriptive tools to investigate the stylistic features of key word, collocation and cluster. He uses two reference corpora of modern fictional texts from the British National Corpus provided (BNC) and a corpus of British prose fiction published during the period 1780 – 1820. He uses the software Wordsmith Tools to extract from the main corpus.

Bujanova (2013) provides a corpus stylistic approach of two famous novels: *Gone with the Wind* by Margaret Mitchell and *A Farewell to Arms* by Ernest Hemingway. The N-Gram tool of AntConc software is applied to the two novels. The extracted N-grams "are outlined in the descending frequency and discussed in reference to the character they can be ascribed to" (Bujanova, 2013, p. 2). Bujanova (2013) adopts Mahlberg's (2012) corpus stylistic analysis of Dicknes' fiction and "Stubbs and Barth's analysis of recurrent phrases functioning as text-type discriminators" (p. 4).

Faiz et al (2018) analyze the category of adjectives in the novel *How to Get Filthy Rich in Rising Asia*. He adopts Leech and Shorts' model (1981) to trace the author's choice of adjectives and their functions. The investigation of the functions of the different types of adjectives in the target novel "throw ample light on the life, culture, economic scenario and love and gender relations construed in the text through adjectives". (Faiz et al 2018, p. 328). Moreover, Spencer (2011) explores the text world of Lovecraft's stories that focuses on the emergence of semantic prosodies via keywords in context, collocation and n-grams.

Hasab El-Nabi (2018) adopts a corpus-based approach to uncover the linguistic features that signal Alice Munro as a feminist writer.

Hasab El-Nabi (2018) compiles a corpus of Munro's thirteen short stories to be tested against a larger corpus of male short stories. Halliday's Systemic Functional Grammar forms the linguistic framework of the study to reveal the lexicogrammatical features that characterize Munro's writing. Hasab El-Nabi (2018) draws a quantitative and qualitative analysis using two software tools: AntConc and TACCO for a grammatical analysis of the cohesive device of reiteration. It is worth mentioning that the current study differs from that of Hasab El-Nabi's (2018) in the linguistic framework, methodology and approach. Whereas Hasab El-Nabi's (2018) thesis analyzes the corpus of the study in terms of Halliday's three metafunctions, this study uses an eclectic linguistic framework based on models from Mahlberg (2007), Sinclair (2004) and Leech and Short (2007). There is another major difference between the two studies which is the current study focuses on lexical cohesion, whereas, Hasab El Nabi (2018) focuses on grammatical cohesion. Moreover, this study deals with corpus tools that differ from the ones used by Hasab El-Nabi (2018). In addition to that, the current study deals with a different literary register and for different objectives. However, the two studies are similar in using the same software used by many researchers for enabling quantative analysis. Most of the above mentioned studies focus on one or two descriptive features and their textual interpretation by analyzing key words, collocation, extended lexical phrases. This study tries to bridge a gap by studying more than one descriptive feature and connecting them to recurrent themes and characters' traits.

3.2. Review of Studies on Young Adult Fiction:

Many studies have analyzed young adult fiction for register and pedagogical aspects tracing the impact of how reading this genre helps in the growth and development of vocabulary. For example, Nelson (2016) compiles a corpus of one million words of Young adult fiction and non -fiction text drawn from the corpus of Contemporary American English database (COCA). Nelson (2016)

draws a comparative study between young adult fiction, children fiction and adult fiction to see if there are any linguistic differences. Nelson (2016) analyzes the corpus using Key word tool to search for the most frequent words in young adult fiction "as opposed to children's fiction and adult fiction" (Nelson, 2016, p. 15). Nelson (2016) finds a difference between children's fiction and young adults' fiction regarding the use of expletives and words referring to body parts and family relationships. His research findings show that young adult fiction "served as a sort of linguistic bridge between children and adult literature (Nelson, 2016, p. 32). For example, expletives are not found in children's fiction; whereas, they prevail in YAF and A F (Nelson, 2016, p. 32). In contrast, words denoting parents are much more common in children literature than in the adult literature. Young adult fiction falls in the middle between the two (Nelson, 2016, p. 15). Moreover, the researcher points out some linguistic features that are more unique to Young adult fiction such as the use of first person pronoun for first person narration. Also, the prevalent use of body parts is a salient characteristic of Young adult fiction (Nelson, 2016, p. 32). Conversely, words referring to animals which are quite frequent in children literature decrease in YA literature (Nelson, 2016, p. 32).

Dror (2014) adopts an ecocriticism approach to dystopian young adult's fiction. He examines the environmental role in shaping young adult's personality. He focuses on the interlink between nature and technology in three dystopian novels: *Life as We Knew It* by Susan Beth Pfeiffer, *The Hunger Games* by Susanne Collins, and *The Knife of Never Letting Go* by Patrick Ness. He explores the ways in which these three texts represent the natural world and technology and attempts to uncover the relationship between the young adult, nature and technology. (Dror, 2014, p. ii)

Al-Pachachi and Naser (2016) draw a comparative study between adult drama and children drama regarding the use of cohesive devices. Their analysis reveals the cohesive device of reference as

the most common in dramatic texts for adults and children (p. 28). On the other hand, ellipsis is more recurrent in adult drama than in children drama. They reach the findings that dramatists' frequent use of cohesive devices communicates effectively their ideas. They also find out that "there is a relation between the age of the readers and the use of cohesive devices as adults are more able to handle certain devices rather than children" (Al-Pachachi & Nase, 2016, p. 28).

Moini and Kheirkhah (2016) study cohesion in children fiction pointing out similarities and differences between adult literature which they call regular novel and children literature with respect to the frequency of lexical cohesive markers and conjunctions. Two novels constitute the corpus of the study: The child novel is *Danny the champion of the World* by Roald Dahl, and the novel selected as the regular fiction is *Coming up for Air* written by George Orwell. Their research findings show that there is a significant dissimilarity between children and adult fiction regarding lexical cohesion devices as well as conjunctions" (Moini and Kheirkhah, 2016, p. 12).

Humanis (2016) also examines the collocation patterns in the novels: I*nsurgent* and *The Fault in Our Stars.* His main concern is finding out the semantic types of adjective in adjective- noun collocations in the novels. He concludes that adjectives depend on personal taste and opinion. He also believes that general adjectives are commonly employed to denote feelings, ideas and experiences.

This study is different from the aforementioned studies in using other descriptive corpus tools that are not dealt with extensively and that the researcher believes that they are more convenient for the purpose of the study. Most studies that have examined lexical devices in literary texts have focused on contrastive analysis between texts of two different genres or registers. Moreover, none of these studies has focused on analyzing in depth a specific register. Hence, this study is projected to fill a gap by focusing on

how the selected lexical semantic devices aid production and interpretation of a particular genre.

4.   Methodology:

This study adopts a qualitative and quantative approach to the analysis of dystopian fiction using the AntConc software (2018) version 3.5.7. AntConc software is chosen for analysis because it is free and user-friendly. Moreover, it provides the researcher with all the descriptive tools that validate the quantification of the corpus such as the Word List tool, Key Word List tool, Collocation tool, N-Grams tool and Concordance tools. While the quantative analysis provides the statistics related to the frequency of collocation patterns and phrasal patterns in the corpus, the qualitative analysis discusses how this can be related to interpretative issues. The corpus of the study is quantitatively and qualitatively analyzed based on:

1. Extracting a word list of the highest frequent words
2. Choosing the highest frequent Nouns as nodes
3. Categorizing the node words according to their semantic preference
4. finding recurrent collocates to the nodes
5. Examining their semantic prosodies
6. Investigating repeated lexical chains

The employed model draws on Leech and Shorts' (2007) checklist for the analysis of the different types of adjectives, Mahlberg's (2007) for the analysis of lexical bundles, Sinclair's (2004) for the analysis of collocation, semantic prosody and semantic preference.

4.2.   Corpus Collection:

The chosen novels cover a period of five years from 2008 to 2013. They are written as trilogies with a different title assigned to each series. They are all categorized in terms of genre as dystopian fiction because they all share the common theme of man-superpower struggle for survival. They are selected according to their rates on the Goodreads list which is "derived from votes from a

large reading community, and reflects which books are most popular " (Nelson, 2016, p. 13).

The corpus comprises the following trilogies with the original word count of each book:

1. The Hunger Games: by Suzanne Collins

   Book 1. (100 000) The Hunger Games, published 2008
   Book 2 (102 000) Catching Fire, published 2009
   Book 3 (101 000) Mockingjay, published 2010
2. The Maze Runner by James Dashner

   Book1 (103 000) The Maze Runner, published 2009
   Book 2 (98 000) The Scorch Trials, published 2010
   Book 3 (88 000) The Death Cure, published 2011
3. Divergent by Veronica Roth
   Book 1 (111 000) Divergent, published 2011
   Book 2 (108 000) Insurgent, published, 2012
   Book 3 (112 000) Allegiant, published 2013
4. Delirium by Lauren Oliver
   Book 1 (115 000) Delirium, published 2011
   Book 2 (94 000) Pandemonium, published 2012
   Book 3 (99 000) Requiem, published 2013

The actual total number of words for the whole corpus is 21,718 of word type and 1,200,3105 word tokens. A stop word list has been pre downloaded into the software to remove all function words that may carry no stylistic value for the purpose of the study. Hence, the total number of words after downloading the stop word list is 21,229 of word type and 449,743 of word tokens.

The corpus is manually downloaded and saved as individual word files, then converted into plain text files to be machine readable. The corpus went through steps of refining by removing images, prefixes, bibliographies so as only the plain text is left. Then the corpus is downloaded into the AntConc software for Word List tool,

Collocation tool, N-Grams tool and Concordance tool. A quantitative analysis of the frequency of occurrences of all content words is done first after removing all function words from the corpus followed by a qualitative analysis of the stylistic significance of the statistical figures for extracting recurrent themes and personal traits of the characters.

5. The Analysis

The first 200 words are generated via the Word List tool of AntConc version 3.5.7 (2018). Then they are searched for the most frequent content words (Figure 1).

Figure 1
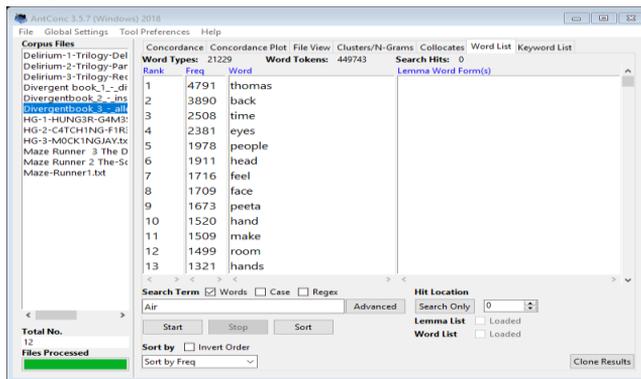
Figure 1. Word List of Content Words

The search results show that the grammatical category of nouns achieves the highest frequency. Leech and Short (2007) classify nouns into general nouns, concrete nouns, abstract nouns, collective nouns. All these types are picked up from the list. However, proper nouns of characters' names are excluded because they are expected to be frequent and also because they carry no significance in relation to the main purpose of this study. This leaves a space to more significant words that may carry more interpretative value. If the same noun occurs twice, once as singular and once as plural, the first occurrence is dealt with. The most frequent nouns are examined for their frequent collocates, semantic preference and semantic prosody.

Table 1

	Rank	Frequency	Keyword
1	3	2508	time
2	4	2381	eyes
3	5	1978	people
4	6	1911	head
5	8	1709	face
6	10	1520	hand
7	11	1499	room
8	13	1321	hands
9	15	1249	voice
10	16	1230	door
11	19	1064	feet
12	26	960	man
13	28	939	things
14	31	870	hair
15	33	868	body
16	34	858	mother
17	35	855	water
18	37	839	day
19	38	833	ground
20	39	828	place
21	42	771	arms
22	43	770	air
23	46	751	wall
24	47	746	life
25	51	716	words
26	54	708	mouth
27	56	703	arm
28	57	699	moment
29	58	696	night
30	60	682	light
31	61	677	mind
32	65	652	gale
33	75	608	girl
34	77	602	shoulder
35	79	589	sound

	Rank	Frequency	Keyword
36	81	582	city
37	83	575	floor
38	84	573	end
39	85	573	raven
40	87	568	district
41	94	542	fingers
42	96	541	world
43	97	528	chest
44	98	522	boy
45	102	504	gun
46	104	501	word
47	105	498	woman
48	107	496	pain
49	109	488	skin
50	110	487	fire
51	115	474	blood
52	117	471	home
53	121	464	heart
54	122	462	father
55	126	457	metal
56	132	444	table
57	133	443	abnegation
58	148	422	house
59	149	420	days
60	150	420	throat
61	151	419	bed
62	152	419	group
63	153	414	point
64	154	413	walls
65	157	406	snow
66	158	405	idea
67	161	403	shirt
68	163	399	food
69	167	390	sounds
70	168	389	glass
71	171	386	years
72	173	383	building

	Rank	Frequency	Keyword
73	177	375	friends
74	180	370	love
75	182	369	neck
76	187	363	maze
77	191	357	stomach
78	192	355	edge
79	193	355	morning
80	195	354	sky
81	198	351	death

Table 1. List of Nouns Based on AntConc Word List Tool

The strength of association between nouns as nodes and its collocates are measured with the MI statistical tool of AntConc. These semantic notions are combined "to provide semantic profiles of words, which show the frequent and characteristic collocates of node words, and make explicit the semantic relations between the collocates." (Stubbs 1995, p.1). The analysis of nouns also fits into Biber's (2011) argument that collocates with similar semantic prosody leads to evaluative meaning of the collocates as good or bad (p. 18). By investigating the semantic preference of nouns, it is found that they can be categorized according to two main groups: nouns referring to different parts of human body and nouns referring to elements of nature. The other nouns are considered miscellaneous as they refer to other categories of different semantic domains. These key noun groups are investigated for their collocation patterns and phrasal structure. In addition to that, their semantic prosodies are also studied for their interpretative and attitudinal value.

Korte (1997) uses the term body language to "refer to the forms of non-verbal behaviour exhibited by characters within the fictional situation" (p. 35). This nonverbal behavior whether intentional or non-intentional, accompanied by an interaction or independent of speech, has communicative and interpretative value in the literary situation (pp.35-36). Korte (1997) argues that body language can express characters' feelings "much more precisely than a verbal

expression can" (p.40). The reason for that is nonverbal behavior tend to be less controllable than verbal behavior. Korte (1997) categorizes body language into modal class and functional class. Modal class refers to 'kinesics' as shown in body movements, facial expressions, eye behavior, and other physical reflection (p. 38). The functional class of body language, on the other hand, is represented by externalizes which "which convey information about a character apart from his or her temporary emotions: relatively stable mental conditions (such as psychopathological states, attitudes, opinions, values, personality traits), but also mental and intellectual activities and conditions" (Korte, 1997, p 41).

Korte 's (1997) classification of body language proves helpful to the analysis; specially the modal one. The frequent use of nouns referring to body parts is related to the fact that characters' feelings are vividly more expressed non-verbally. The excessive use of body language in dystopian fiction serves the function of exposing the characters' feelings, mental and psychological state. It is also used to aid physical description. For example:

"I sneak a look at my reflection when she isn't paying attention— not for the sake of vanity, but out of curiosity. A lot can happen to a person's appearance in three months. In my reflection, I see a narrow face, wide, round eyes, and a long, thin nose—I still look like a little girl, though sometime in the last few months I turned sixteen. The other factions celebrate birthdays, but we don't. It would be self- indulgent.

"There," she says when she pins the knot in place. Her eyes catch mine in the mirror. It is too late to look away, but instead of scolding me, she smiles at our reflection. I frown a little. Why doesn't she reprimand me for staring at myself?" (Divergent)

This extract from *Divergent* abounds in vivid physical description which gives information about the characters. The following section presents a detailed analysis of the different functions of body language as reflected in the frequent occurrence of nouns of body

parts as thematic and textual aspects that characterizes dystopian fiction.

5.1.   Collocation and cluster chains of body parts:

The most remarkable feature in the corpus is the high frequency of nouns referring to body parts and nature. The significance of this lies in the fact that this may suggest that dystopian fiction cares for physical description. The first five nouns referring to body parts are analyzed in context with the accompanying collocates. Table 2 represents the first category of nouns referring to body parts arranged in terms of highest frequency:

Table 2

Rank	Frequency	Word
4	2381	eyes
6	1911	head
8	1709	face
10	1520	hand
15	1249	voice
19	1064	feet
31	870	hair
33	868	body
42	703	arm
54	708	mouth
61	677	mind
77	602	shoulder
94	542	fingers
97	528	chest
109	488	skin
121	464	heart
150	420	throat
182	396	neck
191	357	stomach

Table 2. Nouns of Body Parts

The first 5 nouns referring to body parts are analyzed as nodes together with a span of three collocates to the left. By examining the concordance lines which provide the textual evidence for these body

parts, it appears that they are used to describe either a physical appearance, a body reflect, a romantic scene or aggression. This may appeal to young adults since this is the age of puberty where young adults start developing and growing some physical symptoms and become mentally and psychologically obsessed with their body growth and changes.

5.1.1. The Collocation Patterns and lexical chains of Eyes:

The lexical item "eyes" occupies rank 4 in the word list and occurs 2381 times. This high frequency of this body part is justified by the fact that this lexical item plays a vital role in facial expression. It is essential in expressing emotions and feelings. *Eyes* are also most of the time under supervision by being observed and inspected because they reflect the emotional and mental state of a character. Table 3 shows the collocates of the node word *Eyes.*

According to Leech and Shorts' checklist of adjectives, the adjectives that collocate with the noun "eyes" refer to physical and visual aspects. They are grouped into a) attributive adjectives referring to color and shape of the eyes (such as dark, blue green brown, rounded), b) evaluative and emotive adjectives such as terrified, sleepy, prying, puffy, maddened, fiery. Use of colors to collocate with eyes aids visibility (Sinclair, 2004, p. 33). The collocation patterns of the node *eyes* vary from being normal or usual collocation to unusual/deviant ones at different linguistic levels.

The following Table presents some of the adjectives that collocate with "eyes" from the first highest frequent 100 words. They are categorized according to Leech and Shorts' (2007) classification of adjectives presented in Table 3:

Table 3

Attributive Adjectives of colors (Physical)	Attributive Adjectives of shape (physical)	Adjective of emotive/evaluative attributes
Dark, blue, green, brown, gray, black, yellow, amber, lawny	Wide, round, narrowed, enormous, puffy, angular, pale, waxen, watery, blinking, whimpering, wandering, sparkling, stiffly	Wretched, worried, wild, terrified, suspicious, dead, prying, maddened, glowing, wretched, watery, tired, soulless, stilled, smoldering, blood shot, stinging

Table 3. Types of Adjectives based on Leech and Shorts' Classification (2007)

Figure 2
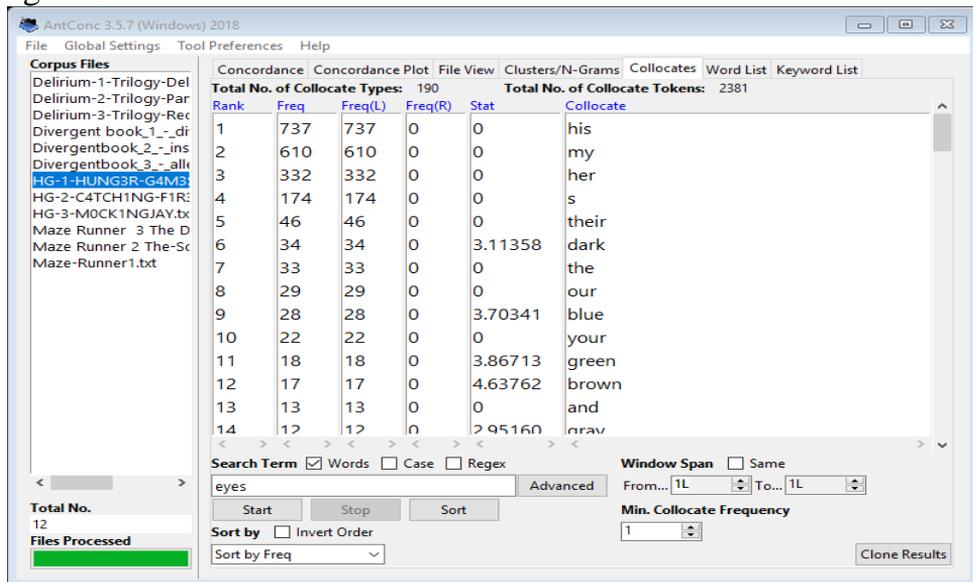
Figure 2. Collocates of Eyes

Eyes are collocated 190 times using a collocation span of one left and one right. These 190 collocates are adjectives with physical, evaluative, visual and emotive attribution according to Leech and Short's (2007) classification. The corpus abounds in a variety of the

collocates of *Eyes* that reflect the mental and psychological state of the characters.

According to figure 2, the lexical item *eyes* demonstrates a strong colligation with possessive pronouns and a strong collocation with adjectives of colors. To the right of *eyes* there is a clear semantic preference to motion verbs such as *closed*, *bulged*, *fell*, *stood*, *show*, *squint* which have the semantic prosody of expressing feeling. By investigating the collocates of *eyes*, we find that the left of *eyes* is occupied by possessive pronouns. "His" occurs 737 times, "my" 610 times, "her" 332 times. Possessive pronouns are very frequent colligates at N1. This shows that possessive pronouns, which are a grammatical category, are essential constituents in the lexical chain with *eyes*.

Adjectives are the second most frequent strong collocates at N2 occurring next to the grammatical category of pronouns and before *eyes*. Thus, collocation and colligation are mixed up together as collocates of *eyes*. Though the lexical item *eyes* is a concrete noun, it is collocated with abstract adjectives with negative semantic prosody. They all occur to the left of the node using the MI measurement tool to test their strength. The variety of the use of adjectives as collocates of *eyes* is rich and significant in the corpus of this research. Adjectives help create a visual image of the events and stimulate imagination on the part of the reader, who in our case is a young adult. They also create a vivid visual image of the characters; thus, leading the reader to visualize them. They appeal to the senses of the reader giving him the joy of visualizing events and characters and making them tangible. "The presentation of physical, psychological, visual, auditory, color and emotive attributes via adjectives has been a source of creating vigour in the novel" (Faiz et al, 2018, p. 332). "The use of color, visual and auditory adjectives also supplement the physical and emotive attributes in order to establish a very vivid picture of the characters, both major as well as the minor" (Faizet al, 2018, p. 332). They help the reader reconstruct a perceptual image of the character

which at the end help him reach an interpretation. Adjectives with emotive attribute tend to evoke a certain emotional response on the part of the reader with either a positive or a negative semantic prosody. Furthermore, the use of adjectives to collocate with nouns helps the reader visualize the events and hence gets more attached and involved with the events as they unravel throughout and with the protagonists of the stories. This is one feature that helps achieve the popularity of young adult fiction. The reader, a young adult, starts growing imagination and so feels obsessed with the elements that address this imagination.

One may hypothesize that the frequent use of *eyes* may presuppose a theme of romance in dystopian fiction. However, by examining the collocation patterns of *eyes*, we find that it implies other themes far from romance in the corpus of this research. This is made clear by the use of deviant and unusual collocation patterns at the semantic level for the adjectives that evoke the attitude of the character. For example, the use of evaluative adjectives such as *wretched*, *wild*, *maddened* and others to collocate with eyes is very much related to the theme of the protagonists fighting against a super power that controls, dominates and watches over the main characters. Therefore, the main characters of dystopian fiction are mostly being monitored by a certain kind of a dominating super power that may also detain and imprison them. This may also justify the frequent use of this particular body part.

Table 4 provides some concordance lines for the emotive and evaluative adjectives which provide the textual evidence upon which interpretation is reached:

Table 4

	Concordance line	Title
1	There's only me looking into the *wretched eyes* of the man from the Nut who asks for one reason.	Mocking Jay
2	*Wild eyes*, hollow cheeks, my hair in a tangled mat.	The Hunger Games
3	Others give me a wide berth, look at me with narrowed, *suspicious eyes*	Insurgent
4	And there are other images too, but they move faster, so I get only impressions of blood and bone and death and cruelty, empty faces, *soulless eyes*, terrified eyes	Insurgent
5	Something flickers across his *bloodshot eyes*	Catching fire
6	Even in his deteriorated state, his *snake eyes* shine bright and cold.	Mocking Jay
7	I wish I could pull the shutters closed, blocking out this moment from the *prying eyes* of Panem.	Hunger Games
8	Seeing them close up—the scratches and blood and bruises, those *maddened eyes*—made Thomas shiver again.	Maze Runner: The Death Cure
9	Their terrified sobs echo in the cement chamber a hundred times, the shrieks of the demons with *glowing eyes*	Insurgent
10	Because at home was my mother with her *dead eyes* and my little sister, with her hollow cheeks and cracked lips.	The Hunger Games

Table 4. Concordance lines of Emotive and Evaluative Adjectives of Eyes

Eyes according to table 4 is collocated with pejorative adjectives: *soulless, wild, maddened* and others. They are considered unusual collocations because they exhibit deviant semantic patterns. They all have negative prosody. The adjective *glowing* in example 9, which tends to have an inherent semantic positive connotation in different contexts, is related to demons in this context. These adjectives serve the function of invoking a certain emotional effect on the part of the reader. It is obvious that "presenting the emotive attributes of characters and events with the help of adjectives creates an emotional effect according to the context in which they are used" (Faiz et al, 2018, p. 332). Adjectives with emotive and evaluative attributes also tend to evoke a certain emotion on the part of the reader.

By investigating the concordance lines of *eyes*, one can see that, it is negatively collocated with negative emotive and evaluative adjectives that serve the functions of creating a certain pictorial image about the characters in concern in the reader's mind. Collocating the eyes with adjectives that have negative connotation may also create a feeling of tension and anticipation of the worst to happen.

Table 5 shows the other collocates of *eyes* that belong to the physical attributes of color:

Table 5

Rank	Frequency Left	Stat	Collocate
6	34	158	dark
9	28	153	blue
11	18	102	green
12	12	115	brown
35	3	2	black
36	3	28	amber
37	2	6	yellow

Table 5. Adjectives of Color

These adjectives are identified by Leech and Short (2007) as visual because they bring visual images of colors to make the scene more visual. The colors "dark" and "blue" achieve the highest frequency. *Dark* collocates with *eyes* 34 times (figure 3), whereas *blue* as a collocate of *eyes* occurs 28 times in the corpus (figure 4).

Figure 3
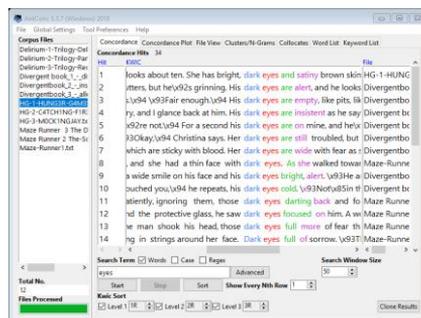
Figure 3. Dark as collocate of Eyes

Figure 4
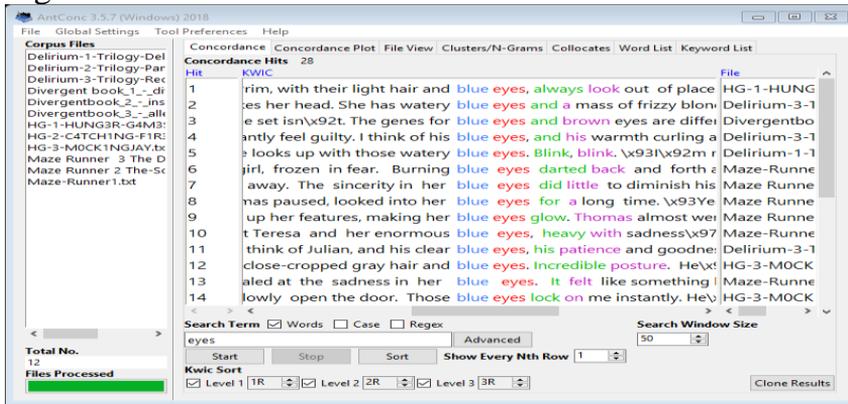
Figure 4. Blue as Collocate of Eyes

According to Hori (2004, p.12) "dark eyes" and "blue eyes" are usual collocations. The following table provides the concordance lines for "dark" and "blue" as collocates of *eyes*:

Table 6

	Textual Evidence	Title
1	Navy Pier," another Dauntless-born initiate calls out. He is tall, with brown skin and *dark eyes*.	Divergent
2	Suddenly I remember him clutching the curtains in his living room, about nine years old, wearing gray, his *dark eyes* shut.	Insurgent
3	She rushes toward me and seizes my hands, which are sticky with blood. Her *dark eyes* are wide with fear as she says, "Are you hurt	Allegiant
4	They erase my face with a layer of pale makeup and draw my features back out. Huge *dark eyes*, full red lips, lashes that throw off bits of light when I blink.	Hunger Game
5	I scan them quickly and then shuffle backward again, getting only an impression of some noses, a few *dark eyes*, the winking of a pair of glasses.	Delirium
6	His *blue eyes*, usually so thoughtful, are now hard and critical, like they are peeling back layer after layer of me and searching each one	Insurgent

	Textual Evidence	Title
7	That's why my mother and Prim, with their light hair and *blue eyes*, always look out of place	The Hunger Games
8	He stares at me. Strangely, despite the rest of his coloring—caramel-and-wheat- blond hair, *blue eyes*.	Pandemonium
9	She has watery *blue eyes* and a mass of frizzy blond hair, which in this light appears to be the same dull gray as the walls.	Requiem
	I think of Julian, and his clear *blue eyes*, his patience and goodness.	Requiem

Table 6. Concordance Lines of Dark and Blue as Collocates of Eyes

By examining the above concordance lines, we can observe that the collocation *dark eyes* is collocated with adjectives associated with negative feelings and sad memories. "*Dark eyes*" is used to describe characters who are scary, huge and vague. In contrast, the color *blue* has the semantic positive prosody of calmness, and purity as it is the color of the sky and sea. It also evokes a feeling of relaxation and tranquility. In the above concordance lines, the collocation *blue eyes* is used as a physical and personal trait. The collocation *blue eyes* is, in many instances, modified by other evaluative adjectives such as *burning* (Hit 6) which stands in strike contrast with the connotation of blue, *enormous* (Hit 10) *clear* (Hit 11) and *watery* (Hit 2) which connote hidden emotion. The collocates of *blue eyes* that appear to the right of the node are either adjectives *darted, heavy, glow* or verbs *look, blink, seemed, shone, show, lock*.

5.2. Lexical clusters of Body Parts:

Examining the clusters of the body parts in the corpus helps the readers get information about characters via their physical description. This proves that body language is essential in communicating information related to characterization. Body clusters are examined using the N-Grams tool of AntConc software. For a phrasal cluster to have statistical significance, it has to occur

at least 10 times. For example, the lexical cluster *eyes on me* achieves a high frequency in the corpus occurring 15 times. It has the semantic prosody of indicating a feeling of romance as it is connected with hidden feelings of love and desire.

Figure 5
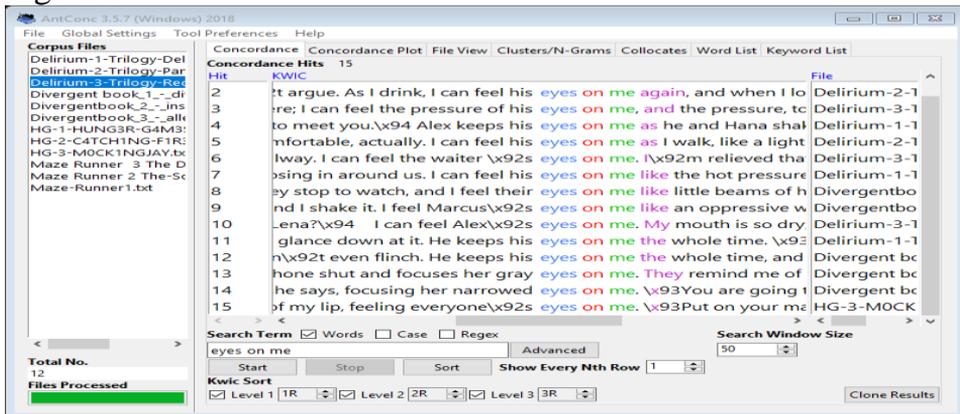
Figure 5. The Lexical Cluster Eyes on Me

My head against (Figure 6) is another frequent lexical cluster which consists of a possessive pronoun followed by the head noun followed by *against*. This lexical cluster occurs 21 times:

Figure 6
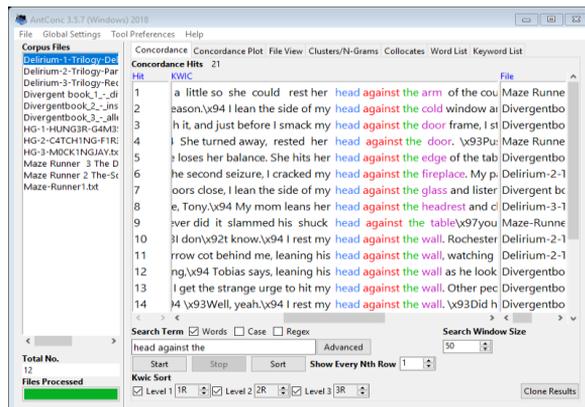
Figure 6. Lexical cluster of (My) Head Against

Table 7

1	Thomas closed his eyes for a second and rested *his head against* the wall.	the Scorch Trial
2	Teresa stretched out a little so she could rest *her head against* the arm of the couch.	the Scorch Trial
3	I smile into my hand, lean *my head against* the window, and let the tears fall in silence My parents did love each other. Enough to forsake plans and factions.	Allegiant
4	I rest *my head against* the wall "Did he think he could rescue you?" Peter snorts a little.	Insurgent
5	I get the strange urge to hit *my head against* the wall.	

Table 7. Concordance Lines of (my) Head Against

The lexical chain (*my*) *head against* is accompanied by another body movement such as closing eyes as in Hit 1 or stretching body in Hit 2, smiling Hit 3 and leaning Hit 4. These body movements are usually associated with recalling past memories reflecting mental and emotional state. This body part cluster acts as a trigger for events to follow. The character's talk is also accompanied by a physical description of the head and other associated body parts. Moreover, they signal the beginning of a dialogue between the characters. Thus, "body Part clusters can provide contextual information that accompanies the description of a situation or activity which is more central to the story" (Mahlberg, 2007, p. 25). These clusters lie at the heart of description to reflect a habitual behavior. Furthermore, it is a fixed textual pattern that appears at the beginning of a sequence of activities, thus, serving narration purposes.

Head against the wall is another lexical chain that occurs 11 times.
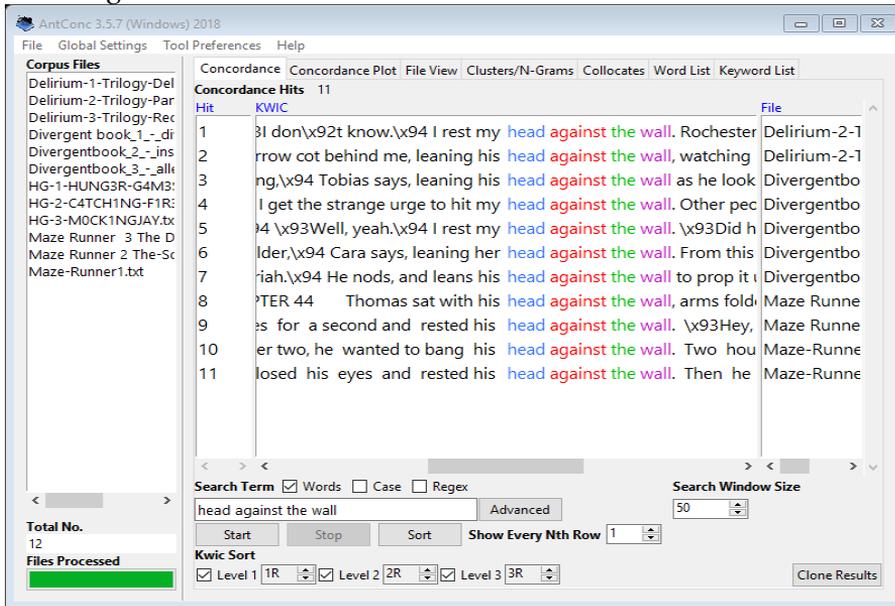
Figure 6. Lexical cluster of Head against the Wall

The following table provides some concordance lines that include the textual evidence for this recurrent lexical chain.

Table 8

1	"I don't know." I rest my *head against the wall*. Rochester, New Hampshire. I must have looped around the northern border when I was lost in the Wilds: I've ended up sixty miles southwest of Portland.	Pandemonium
2	This time I do sit up, whipping around despite the pain. Julian Fineman is sittingon a narrow cot behind me, leaning his *head against the wall*, watching me through heavy-lidded eyes. He is holding a tin cup, which he extends toward me.	Pandemonium
3	Thomas sat with his *head against the wall*, arms folded. Teresa had her legs pulled up under her,	The Death Cure

	sitting so that she faced him. Minho had warned him not to listen to a word she said as they'd walked away.		
4	The first hour passed, and Thomas felt boredom creep in like rats under the door. By hour number two, he wanted to bang his *head against the wall*.	The Maze Runner	

Table 8. Concordance Lines of Head against the Wall

The body part *head* in these examples has a collocation cluster of the grammatical structure of a noun followed by a prepositional phrase and the whole chain functions in Hits 1 and 3 and 4 as the object of the main verb. It describes a physical state followed by a state of physical activity. *Head* is colligated with a possessive pronoun in all hits

"*A hand on*" is another frequent cluster in the corpus. It occurs 41 times.

Figure 7
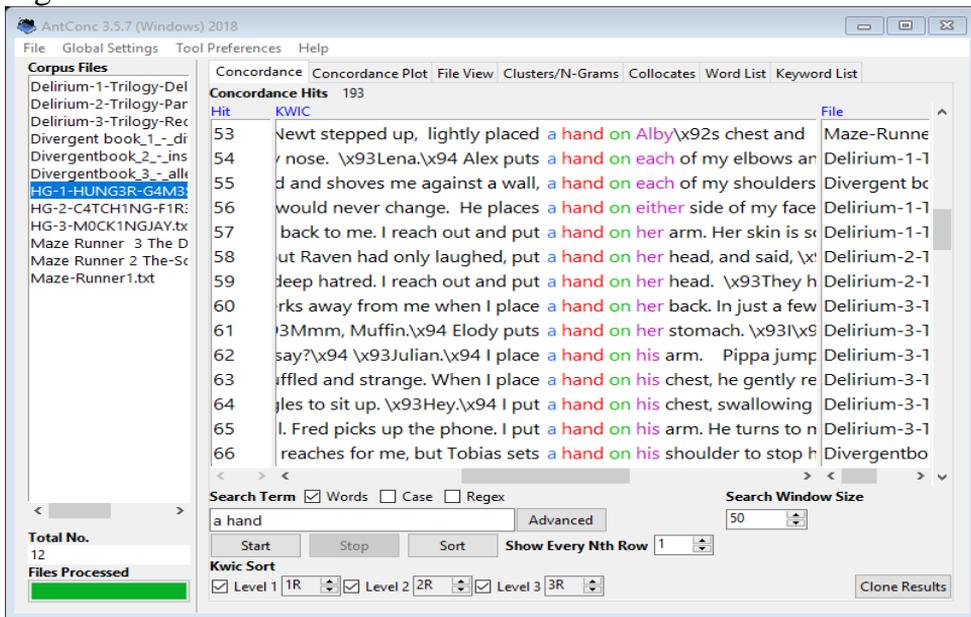
Figure 7. Lexical Cluster of a Hand on Me

Some concordance lines for *a hand on* are presented in the following table:

Table 9

	Concordance lines	Title
1	Killegan places *a hand on* my arm. "Don't worry, dear," she says. "We'll find the perfect dress. That's what this is all about, isn't it?"	Requiem
2	Coral puts *a hand on* my arm and draws me toward a thick triangle of shadow: a recessed doorway, scented with cat urine and cigarette smoke and half-concealed behind a pillared entry. We crouch in the shadows.	Requiem
3	He walks toward his bunk, but at the last second, he whips around and shoves me against a wall, *a hand on* each of my shoulders	Divergent
4		Delirium

Table 9. Concordance Lines of a Hand on Me

This lexical chain *a hand on* has a fixed grammatical pattern of a head noun followed by a prepositional phrase. It contributes to the dynamicity of the action as in Hit 3 which is crammed with short actions that accelerate the events. This creates an atmosphere of suspense which is an essential characteristic of dystopian fiction. The mentioning of the body part *hand* and sometimes in association with other body parts as *arm*, *palm* and *head* provides some contextual information. In Hits 2 and 5 the clusters *a hand on my arm* have the semantic prosody of creating a sense of relief in moments of despair. In Hit 1, the body parts *hand and palm* reflect a psychological and mental state of the character revealing a feeling of deep worry and suspense (the sweat on the palm).

In contrast, *heart* and *lips* which may have the semantic prosody of romance are the least frequent with *heart* occurring (464) times and lips (336) times when compared to other words referring to different body parts such as feet. This, may indicate that romance is not a major theme in dystopian fiction. This goes in contradiction to Nelson's (2016) findings which confirm that there is an excessive use of words of human bodies used to construct a subplot of romance in his corpus of YA as opposed to adult fiction. He indicates that this frequent use of body parts in association with

romance is a major trend in young adult fiction (p. 32). This is, however, not the case according to the findings of the analysis. This may be justified by the fact that dystopian fiction belongs to the science fiction genre where the desire to survive is more valued than being loved.

The frequent use of body parts to imply actions communicates a lot about characters' attitudes and feelings. Eye gaze, hand and head movement imply action and actions "are deemed to be expressive" (Muller et al, 2013, p. 7). Facial expressions, and body gestures " show each other affection, disdain, indifference, concern, gratitude; they challenge or threaten one another; they submit, comply, or defy one another, or they show fear, joy, and so on" (Muller et al, 2013, p. 7).

5.3.   Collocations of Nouns referring to Nature:

The second category of nouns that is higher in frequency contains words referring to nature. The first 500 words are searched for the highest frequent words that are related to nature. This high frequency may reflect the growing interest of young adults in exploring the surrounding nature. As they grow up, young adults develop a relationship with nature. Taken into consideration that the corpus is a dystopian fiction, one expects the theme of how to timid a revolting nature to be prevalent in dystopian fiction. Nature is not harmful to humans; it is a human victim because it is being abused by humans and climate change is the result of human corruption and this is responsible for all the conflicts that lead the characters to be trapped and detained by a dominant authority. Nature is being victimized, yet the superpower leads the main characters to believe that they are the real victims of climate change and these superpowers are protecting them from the danger of nature. In fact, these superpowers have hidden agendas. Most young adult fiction protagonists have insecure relationships with nature (Dror, 2014). In dystopian fiction, young adults are either nature victims or nature savers. Young adults are invited to

contemplate the sinister reality of nature and think of a way to win it back. Collocates of nouns referring to nature are examined through concordancing.

Table 10 below presents a list of nouns denoting nature:

Table 10

Rank	Frequency	Word
188	855	water
191	839	day
192	833	ground
205	770	air
229	696	night
242	652	gale
318	487	fire
377	406	snow
425	355	morning
428	354	sky
463	322	haven
475	312	trees

Table 10. Word List of Nouns referring to nature

5.3.1. Collocations and the lexical chain of Water and Ground:

Water occurs 885 times with total number of collocate types 737 and collocate tokens of 2565. Only collocates that occur at least 10 times are taken into consideration. *Water* is used throughout to refer to natural sources such as rivers and seas, and sometimes in its natural sense as a drinkable liquid:
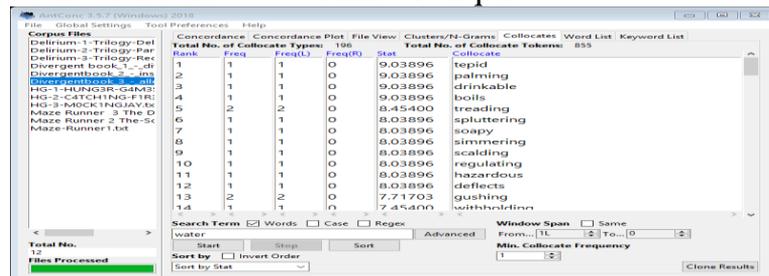
Figure 9. Collocates of water

However, *water* is used in many instances with negative prosody. The following concordance lines provide the textual evidence for the use of *water* as a threatening natural resource:

Table 11

	Textual Evidence	Title
1	I feel like I just walked into a downpour, and all my clothes are heavy with *water*; like I am heavy and awkward and useless.	Divergent
2	In another I was drowning again, this time in the middle of an ocean as the *water* raged around me.	Divergent
3	Wind tore through the air, pelting him with small rocks and drops of *water* that hurt equally.	The Maze Runner
4	In a home, or a brightly lit nursery or office, it might be cheerful; but illuminated only by the patchy fluorescent lights that keep buzzing on and off, and stained with years and years of *water* and handprints and squashed insects and I don't-want-to-know-what, it seems incredibly depressing—like getting a big smile from someone with blackened, rotting teeth.	Delirium
5	There's an archway on our right—a cutout of even blacker darkness—and I squeeze Julian's hand, pulling him back, directing him through it, into another tunnel, a foot or so lower than the one we've been traveling, and this one dotted with puddles of stagnant, *stinking water*.	

Table 11. Concordance lines of Water

Water in these concordance lines is mentioned in association with harsh natural elements. It constitutes a source of threat to the characters in concern. In the above Hits, it is used metaphorically to indicate a difficult situation the character is going through. However, there are instances where *water* is used in its natural sense as a drinkable liquid. It is collocated with food 12 times with the MI strength of 4.79 to refer to water in its natural sense. *Food* and *water* constitutes a cluster of frequent occurrence.

Figure 10
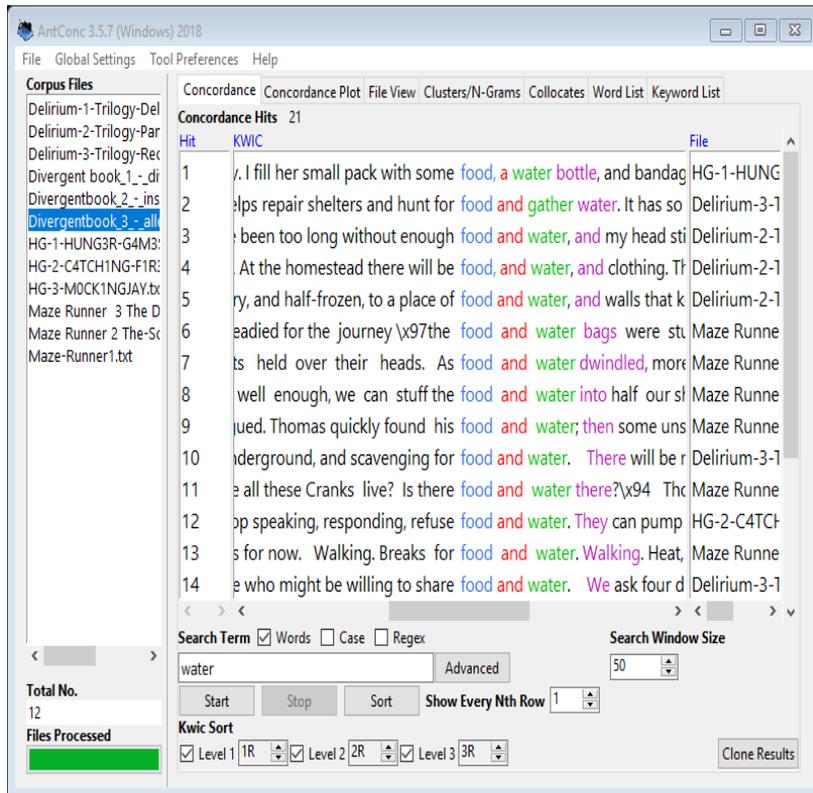
Figure 10. Food as Collocate of Water

Table 12

1	Even though I feel better than I did earlier, I'm still not moving very well. I've been too long without enough *food and water*, and my head still throbs. My left ankle wobbles as I hit the ground, and for a minute I stumble against Julian, bumping my chin on his chest, and his arms tighten around me.	Pandemonium
2	But for now, the future, like the past,	Pandemonium

	means nothing. For now, there is only a homestead built of trash and scraps, at the edge of a broken city, just beyond a towering city dump; and our arrival— hungry, and half-frozen, to a place of *food and water*, and walls that keep out the brutal winds. This, for us, is heaven.	
3	The others don't come. Eventually we are forced to admit that we've been separated from them. Raven says, halfheartedly, that they will no doubt track us down. We need to find somewhere we can safely camp, and someone who might be willing to share *food and water*.	Requiem
4	Today, I go east. It's one of my favorite times of day: that perfect in-between moment when the light has a liquid feel, like a slow pour of syrup. Still, I can't shake loose the knot of unhappiness in my chest. I can't shake loose the idea that the rest of our lives might simply look like this: this running, and hiding, and losing the things we love, and burrowing underground, and scavenging for *food and water*.	Requiem
5	"Found a gap through the mountains,"Minho answered. "Had to fight through some Cranks amping in a cave, but other than that, no problems. *Food and water*'s almost out, though. And my feet hurt.	The Scorch Trials

Table 12. Concordance Lines of Water and Food

The lexical chain food and water occurs 15 times in almost similar contexts. The characters are in the middle of somewhere suffering

from food and water shortage. A distressful feeling is felt due to the fact that their physical needs are not met. This aggravates the harsh circumstances they go through.

Moreover, *water* is collocated with adjectives denoting physical features such as color, temperature, size (black, blood stained, cold, hot, pool) with no statistical significance.

The node "ground" occurs 833 with the total number of collocate types 63 and collocate tokens 833. It is collocated with theses adjectives: solid (MI 7.058), flat (5.39), desert (8.098), uneven (8.56774), unstable (7.77).

Figure 11
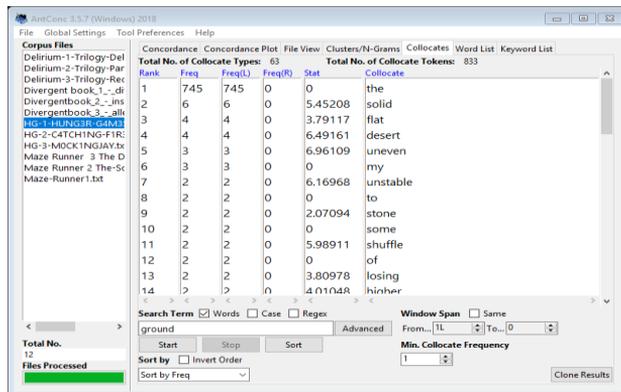
Figure 11. Collocates of Ground

Table 13 includes some concordance lines for "ground":

1	A weightless moment, and then my feet slam into *solid ground* and pain prickles through my shins.	Divergent
2	Soon the rotating square had fully flipped, and where once had been *desert ground* now lay a section of black material, with an odd object sitting on top of it.	The Scoorch
3	At the end of each hallway is a blue lamp,	Divergent

	but between them it's dark, and I have to be careful not to stumble over *uneven ground*.	
4	*The rain-softened ground* made it hard to get good traction; Thomas slipped twice, fell down once.	The Scorch
5	On three we launch off the train car. A weightless moment, and then my feet slam into *solid ground* and pain prickles through my shins.	Divergent

Table 13. Concordance lines of Ground

These adjectives describe the physical nature of a ground with negative prosody because they connote a harsh and deserted aspect of nature.

Air is another noun referring to nature that occurs frequently in the corpus. It occurs 770 times with the total number of collocate types 117 and collocate tokens 770. It is frequently collocated in the corpus with the physical attribute adjectives "cool" with MI of (8.055) "cold" (6.44), "thin" (7.12), open (4.69) and "fresh" (7.50). These are usual collocation patterns with expected collocates of the node.

Figure 12
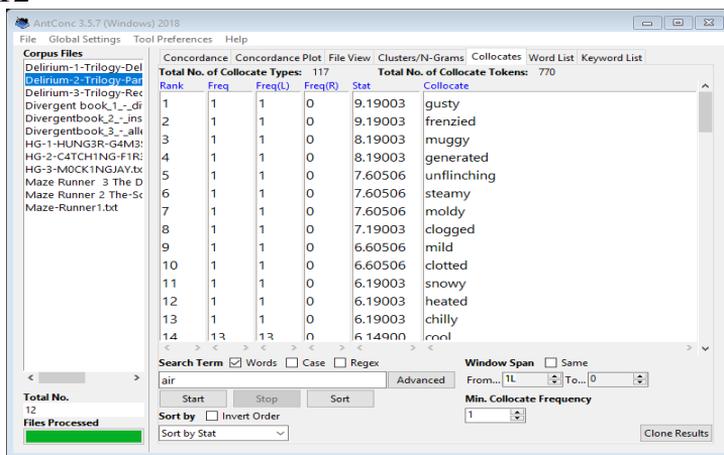
Figure 12. Collocates of Air

Table 14

	Concordance Lines	Title
1	The ramp door started cranking open with the squeal of hinges while they were still a hundred feet up, and *cool air* blasted inside.	The Death Cure
2	Just then a shadow zooms past me—so sudden and startling I nearly slip backward. For a moment I feel the terror of free fall —the tipping, the *cold air* behind me— but at the last second I manage to right myself.	Pandemonium
3	Then I hesitate a moment, thinking of another day in this woods, when Gale and I watched a hovercraft appear out of *thin air* and capture two escapees from the Capitol.	Catching Fire
4	I pound up the stairs and throw myself outside, into the *open air* and the warm afternoon and the throaty sound of the woods opening up to spring.	Requiem
5	Otherwise you may never know hell, but you will not find heaven, either. You will not know **fresh air** and flying.	Requiem

Table 14. Concordance lines of Air

The attributive adjectives that collocate with the lexical item "air" have the function of easing a tense atmosphere.

The lexical chains of *air* and *ground* are found to be of non-significant statistical value since they achieve minimal frequency. For a lexical set to be considered as a chain, it has to appear at least 10 times. Therefore, it is excluded from the analysis.

The frequent use of adjectives with negative semantic prosody as collocates of the discussed nodes shows that nature for the young protagonist is either a threat or a safeguard. The theme that nature has been subjected and manipulated by technology

6. Findings and Conclusion:

The first hypothesis of the study about how corpus linguistics validates literary analysis is proved to be valid for several reasons. First, the study shows how corpus linguistics aids the exploration of textual and thematic features that may transcend traditional stylistic analysis. Second, corpus linguistics relies on finding the evidence for whatever interpretation is reached in the text itself. Thus, it has proved helpful in providing the researcher with a list of the highest frequent words and the concordance lines that provide the textual evidence for validating an objective interpretation of some important characteristics of the texts. Corpus linguistics also aids the analysis of important lexical and semantic features such as collocation, semantic prosody and semantic preference and lexical clusters that may be elusive to analysis by traditional stylistics when investigated in large corpora.

The second hypothesis of the research is also tested by conducting a lexical semantic analysis of the most frequent lexis and patterns in the corpus of the study. It has been revealed through the analysis of collocation, semantic prosody, semantic preference, and lexical clusters that dystopian fiction has unique stylistic features on one hand and solid texture on the other. Analyzing collocation patterns in fiction enriches our understanding of the literary text.

In answer of the first and second research questions, the corpus of the study is first searched for content words via the Word List tool of AntConc. Then the grammatical category of nouns is chosen for analysis as they are content words. They are categorized according to their semantic preference into two groups: nouns referring to body parts and nouns referring to nature. This grouping of nouns into two broad categories based on their semantic preference can be indicative of major stylistic features of dystopian fiction. It has been revealed that nouns referring to body parts and nouns referring to nature are of high frequency in the corpus of the study. These nouns

were further investigated in terms of their collocation patterns, semantic prosody and lexical chain.

In answer of the third research question about the types of adjectives that collocate with nouns, attributive and evaluative adjectives are found to be the most frequent adjectives. They are examined as collocates of nouns and their function is identified in the context. Evaluative and emotive adjectives are exploited to show a habitual behavior of characters and to create a specific and peculiar tone that reflects the atmosphere of fictions. The findings of the study revealed the adjective-noun collocation as the most prevailing pattern. Adjectives as noun collocates belong to the attributive and evaluative group according to Leech and Short Checklist. They color the situation with emotions and feelings on one hand and reflect recurrent themes in the corpus of the study on the other. It also shows that dystopian fiction cares for physical description.

In answer to the fourth about the function of body parts cluster, findings revealed that frequent occurrence of nouns referring to different parts of the human body is associated with the important role played by body language in dystopian fiction to identify the characters' physical and habitual traits in the corpus. Moreover, the use of body parts associated with gestures and motion (body language) achieves authenticity of events and situations "by creating a lively or life-like picture" (Mahlberg, Smith & Preston, 2013, p. 41). In general, body language such as body gestures, facial expression, pitch of voice and other paralinguistic features are clues of characters' feelings and emotions. Mahlberg et al (2013) argue in favor of this point that "patterns of body language presentation can be seen as 'key phrases' in narrative fiction, in the sense that body language plays an important role in the creation of characters"(p.41). They help readers recognize fictional characters through their habitual behavior (Mahlberg et al, 2013, p. 41). Consequently, it becomes an essential feature in fantasy fiction because it aids a vivid description of characters' emotion. Moreover, the frequent occurrence of words referring to body parts may be

justified by the fact that the reader is a teenager who is starting growing interest in the physical change in his body. Thus, it is counted as an appealing characteristic of the genre to the target reader. Shortly, the excessive use of body language aids visualization of characters and events. Lexical clusters of body parts, on the other hand, are used "to fulfil local textual functions in the creation of fictional characters" (Mahlberg et al, 2013, p. 41). A contextual analysis of some of these collocation patterns of body parts shows that they serve the function of recalling memories, initiating a series of events or expressing the inner thoughts and emotions of a character. Moreover, most of the collocation patterns used for body parts exhibit deviant semantic patterns.

6.1. Suggestions for Further Research:

The interplay between language and literature is a fruitful area for further studies by exploring other aspects of literary texts from linguistic perspectives. First, the scope of the study can be shifted from corpus stylistic approach to a cognitive corpus stylistics by using models that belong to cognitive linguistics such as schematic structure, mental spaces and conceptual metaphors which emphasize the ability of the reader to construct meaning cognitively. Accordingly, research in cognitive stylistics may exploit tools such as questionnaires and interviews to elicit readers' responses. Second, researchers can implement other techniques, theories and tools offered by linguistics to be experimented on literary texts of different genres. Third, the eclectic model of the current research may be further applied to non-literary texts for different objectives. One of the limitations of this study is that there is no reference corpus to compare to the main corpus for keyness value. Hence, other studies may extend this research by having a reference corpus of either adult literature or children literature for the key word list. Moreover, the study of the semantic domains of highest frequent words using other computational software programs such as the W matrix can be more helpful in locating recurrent themes. Further

studies may have a pedagogical scope by studying how collocation can be used in EFL classes. In addition to that, the variation in register between dystopian fiction and young adult fiction can be examined for similarities and differences.

References

I.   Primary References

Collins, Suzanne (2008).*The Hunger Games*. Scholastic, Inc.
Collins, Suzanne (2009). *Catching Fire*. Scholastic, Inc.,
Collins, Suzanne (2010). *The Mockingjay*. Scholastic, Inc.
James Dashner (2009). The Maze Runner. Delacorte Press
James Dashner (2010). The Scorch Trials. Delacorte Press
James Dashner (2011). The Death Cure. Delacorte Press
Lauren Oliver (2011). *Delirium*. Harper Collins.
Lauren Oliver (2012). *Pandemonium* HarperCollins
Lauren Oliver (2013). *Requiem* HarperCollins
Roth, Veronica (2011). *Divergent*. Harper Collins.
Roth, Veronica (2012). *Insurgent*. Harper Collins.
Roth, Veronica (2013). Allegiant. Harper Collins.

II. Secondary References

Al-Pachachi, A., & Nase, A. (2016). The Use of Cohesive Devices in
    English Drama Texts for Adults and Children. *Angloamericanae
    Journal, 1*(1), 28-60. doi: 10.5281/zenodo.1069659
Balossi, G. (2014). *Corpus Linguistic Approach to Literary Language
    and Characterization : Virginia Woolf's.* John Benjamins
    Publishing   Company.   Retrieved   from
    http://ebookcentral.proquest.com
Begagić, M. (2013). Semantic preference and semantic prosody of the
    collocation make sense. *Jezikoslovlje, 14* (2-3), 403-416.
Biber, D. (2011). Corpus linguistics and the study of literature: Back to
    the future? *Scientific Study of Literature, 1* (1), 15-23
Bujanova, K. (2013). *A Corpus-Stylistic Analysis of Mitchell's Gone with
    the Wind and Hemingway's A Farewell to Arms.* (Unpublished
    M.A. Thesis). University of Oslo.
 Cerban, M. (2010). Lexical Cohesion: Aspects of Collocation using
    Halliday and Hasan's Systemic model of cohesion. *Translation
    Studies: Retrospective and Prospective Views*. Retrieved from
    https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2799260

Conrad, S., & Biber, D. (2004). The frequency and use of lexical bundles in conversation and academic prose'. In W. T. Mahlberg (Ed.), *The Corpus Approach to Lexicography, Thematischer Teil von Lexicographica* (pp. 56–71). Niemeyer: Tübingen.

Dror, S. (2014). *The Ecology of Dystopia: An Ecocritical Analysis of Young Adult Dystopian Texts*. (Unpublished M.A. Thesis). The University of British Columbia, Vancouver.

Faiz, R., Azher, M., Asghar, I., & Jabeen, I. (2018). Stylistic Study of Adjectives in How to Get Filthy Rich in Rising Asia. *International Journal of English Linguistics, 8*(3), 328-338. doi:10.5539/ijel.v8n3p328

Firth, J. (1957). *Papers in Linguistics.* Oxford: Oxford University Press. .

Fischer-Starcke, B. (2010). *Corpus Linguistics in Literary Analysis: Jane Austin and her contemporaries.* London, New York: Continuum International Publishing Group.

Flowerdew, J., & Mahlberg, M.(Eds.). (2009). *Lexical Cohesion and Corpus Linguistics.* Amsterdam / Philadelphia: John Benjamins Publishing Company. doi:https://doi.org/10.1075/bct.17

Giannossa, L. (2012). A Corpus-Based Investigation of Lexical Cohesion In En & It Non-Translated Texts And In It Translated Texts (Unpublished Ph D Dissertation). Kent State University.

Hakim, R. A. (2016). Lexical Cohesion in Fiction Stories With Reference to the Frog Prince and the Bully. Humanist, [Sl]*,* 13(2), 1-7. Retrieved from HYPERLINK "https://ojs.unud.ac.id/index.php/sastra/article/view/17222" \t "_new" https://ojs.unud.ac.id/index.php/sastra/article/view/17222

Halliday, HYPERLINK "https://www.google.com.eg/search?tbo=p&tbm=bks&q=inauthor:%22M.A.K.+Halliday%22" M.A.K. ,& HYPERLINK "https://www.google.com.eg/search?tbo=p&tbm=bks&q=inauthor:%22Ruqaiya+Hasan%22" Hasan , R.(1976) *Cohesion in English*. London: Longman

Hassab El-Nabi, E. (2018). A Corpus-Based Linguistic Study of Contemporary Feminist Short Story Writers: With Special Reference to Alice Munro (Unpublished PhD Dissertation). Faculty of Women. Ain Shams University.

Hoey, M. (1991). *Patterns of Lexis in Text*. New York/Oxford/Toronto: Oxford University Press.

Hu, M. (2015). A semantic prosody analysis of three adjective synonymous pairs in COCA. *Journal of Language and Linguistic Studies, 11*(2), 117-131.

Jobbins, A., & Evett, L. (1998). Text Segmentation Using Reiteration and Collocation. Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics. Montreal, Quebec, Canada. Retrieved from www.aclweb.org/anthology/P98-1100.

Korte, B. (1997). Body Language in Literature. University of Toronto Press.

Leech, G and Short, M. (2007). *Style in Fiction: A Linguistic Introduction to English Fictional Prose* (2nd ed.). London and New York: Longman.

Louw, B. (1993). *Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies.* In M. Baker (Ed.), *Text* and Technology: In Honour of John Sinclair. Amsterdam – Philadelphia: John Benjamins.

Mahlberg, M. (2006). Lexical cohesion: Corpus Linguistic Theory and its Application in English Language Teaching. *International Journal of Corpus Linguistics, 11*(3), 363–383.

Mahlberg, M. (2007). Clusters, key clusters and local textual functions in Dickens. *Corpora, 2*(1), 1-31.

Mahlberg, M. (2009). *Lexical cohesion Corpus linguistic theory and its application in English language teaching.* In J. a. Flowerdew (Ed.), *Lexical Cohesion and Corpus Linguistics.* Amsterdam / Philadelphia: John Benjamins Publishing Company.

Mahlberg, M. (2010). Corpus Linguistics and the Study of NineteenthCentury Fiction. *Journal of Victorian Culture, 15*(2), 292-298.

Mahlberg, M., & McIntyre, D. (2011). A case for corpus stylistics Ian Fleming's Casino Royale. *English Text Construction, 4*(2), 204–227. doi:doi 10.1075/etc.4.2.03mah

Mahlberg, M., Smith, C., & Preston, S. (2013). Phrases in literary contexts Patterns and distributions of suspensions in Dickens's

novels. *International Journal of Corpus Linguistics, 18*(1), 35–56. doi:doi 10.1075/ijcl.18.1.05mah

Malah, Z. (2016). Signals from the Surface of Texts: How Lexical Cohesion Reflects Writers' Tones in Nigerian Newspaper Editorials. *International Journal of Language and Linguistics, 3*(5), 175-196.

McIntyre, D. (2013). *Corpora and Literature.* In C. Chappelle (Ed.), *Wiley Blackwell Encyclopedia of Applied Linguistics*. Wiley Blackwell: Wiley Blackwell Encyclopedia of Applied Linguistics.

McIntyre, D. (2015). Towards an integrated corpus stylistics. *Topics in Linguistics*(16), 59-68. Retrieved from   HYPERLINK "http://dx.doi.org/10.2478/topling-2015-0011" http://dx.doi.org/10.2478/topling-2015-0011

Moini, M., & Kheirkhah, F. (2016). Use of Cohesive Devices in Children and Regular Literature: Conjunctions and Lexical Cohesion. *International Journal of Comparative Literature & Translation Studies,   4*(4),   12-23.   Retrieved   from   URL: http://dx.doi.org/10.7575/aiac.ijclts.v.4n.4p.12

Müller, C., Cienki, A. J., Fricke, E., Ladewig, S. H., McNeill, D., & Teßendorf, S. (Eds.). (2013). *Body- Language-communication: an International Handbook on Multimodality in Human Interaction.* Berlin: De Gruyter Mouton.

Nelson, K. M. (2016). *Lexical Trends in Young Adult Literature: A Corpus Based Approach* (Unpublished M.A. Thesis). Brigham Young University.

O'Halloran, K. (2007). The subconscious in James Joyce's 'Eveline': a corpus stylistic. *Language and Literature, 16*(3), 227-244.

https://en.oxforddictionaries.com/definition/utopia. (n.d.).

Philip, G. (2011). *Colouring Meaning: Collocation and Connotation in Figurative Language.* Amsterdam / Philadelphia: John Benjamins.

Sinclair, J. (2004). *Trust the Text: Language, corpus and discourse.* London and New York: Routledge: Taylor and Francies Gorup.

Sinclair, J. (2010). *Exploring Gothic Fiction: A Corpus-Based Analysis.* Edinburgh: The University of Edinburgh.

Spencer, H. (2011). Semantic Prosody in Literary Analysis: A Corpus based Stylistic Study of H. P. Lovecraft's stories(Unpublished M.A. Thesis). University of Huddersfield. Retrieved from available at http://eprints.hud.ac.uk/id/eprint/12904/

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of trouble with quantitative studies. *Functions of Language, 2*(1), 23–55.

Stubbs, M. (2005). Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature, 14*(1), 5-24.

Traugott, E., & Pratt, M. L. . (1980). *Linguistics for students of Literature*. New York: Harcourt Brace Jovanovich.

Utopia (n.d.) In Oxford Dictionary. Retrieved from https://en.oxforddictionaries.com/definition/utopia

Utopia & Dystopia (n.d.). Definition, Characteristics & Examples. Retrieved from https://study.com

Utopia and Dystopia. (2018). Dystopian Elements and Characteristics - Basic Building Blocks of Dystopia. Retrieved from www.utopiaanddystopia.com

Wijitsopon, R. (2013). A Corpus-Based Study of the Style in Jane Austen's Novels. *Manusya: Journal of Humanities Regular, 16*(1), 41-64.

Wynne, M. (2006). Stylistics: Corpus Approaches in K.Brown (Ed.). *The Encyclopaedia of Language and Linguistics (pp. 223–26).* Oxford: Elsevier.

Wynne, M., & Prytz, B . (2008). Corpus Approaches to the Language of Literature. T*he International Conference of Corpus Linguistics in Practical Applications*. Viterbo, Italy

Xiao, R., & Mcenery, T. (2006). Collocation, Semantic Prosody, and Near Synonymy: A Cross-Linguistic Perspective. *Applied Linguistics, 27*(1), 103–129 . doi:10.1093/applin/ami045

Young, M. (2011, October, 23). Why is dystopia so appealing to young adults? Retrieved from https://www.theguardian.com/books/2011/oct/23/dystopian-fiction