Differences in Frequencies between Linking Verbs and Relative Pronouns in Written Language

Thanaa Alhabuobi *thamar_5@hotmail.com*
Arts & Humanities/ Language & Translation
Taibah University
Almadinah Almunawarrah/ 344/ Saudi Arabia

Abstract

This study was conducted primarily to investigate two areas of linguistic analysis. The study aimed to analyze two types of associations: lexical-grammatical associations and grammatical-grammatical associations. The method used in this study was descriptive and analytic. With regard to lexical-grammatical associations, four linking verbs were selected to explore differences among them with respect to their occurrences in the selected texts. Concerning grammatical-grammatical associations, two relative pronouns were selected to identify differences among them with respect to their occurrences in the selected texts. The pronouns were specifically selected in terms of their use as a subject in essential (restrictive) adjective clauses. For purposes of the current research, seventeen articles were selected from Internet databases. The topics of all articles involved health-life problems. The Simple Concordancer Program was applied to find collocations and frequencies of the target words. The study was delimited to four linking verbs: *be, seem, look,* and *feel*. It was also delimited to two relative pronouns used as a subject in restrictive adjective clauses. The findings of the study revealed no significant differences between occurrences of the linking verbs. However, the verb (be) occurred relatively more frequently than (*seem, look*, and *feel*). On the other hand, the verb (look) was the least frequent verb among all verbs. The findings also showed significant differences between occurrences of the subject relative pronouns. More specifically, the pronoun (who) occurred more frequently than the pronoun (that).

Keywords: Linguistics, Computer, Concordances, Linking Verbs, Pronouns.

1. INTRODUCTION

There have been an enormous number of language studies that explore the nature of linguistic data. Perhaps, this work was to explore why certain patterns are common while others are not. However, most studies emphasized the structure of language. In addition, these studies handled a relatively limited number of corpora. Later, the idea of investigating linguistic data and what relates to its associations became different. The impact of technology and the use of the extended corpus-linguistics led to a tangible change in areas of linguistic research. Given these factors, many new aspects appeared and became plausible and more substantial than the old ones. The relationship between genre and the use of language patterns became worthy of being studied as a means of exploring which patterns are more common in fiction work than in non-fiction work. Another example involves identifying differences between written forms and spoken forms in terms of grammar, lexicon, pragmatics, and semantics. The invention of the computer also contributed to this matter; previously, this labor usually took a lot of time and did not produce reasonably accurate results. Nowadays, it has become faster, easier, and more accurate with the use of computational linguistics.

1.1 Statement of the Problem

Investigating various associations that occur in language structures became a modern approach to meet the new trend in linguistic research. These associations are of different types and present a rich field upon which linguists can explore the characteristics of both language use and

language structure (Biber, et al., 1998). The impact of computer use opens the door wide for corpus linguistics to be an applicable approach and an ample device for identifying and analyzing various associations across language structures. Consequently, the analyses of those associations are needed to identify different language uses, generalize rules, and explain the nature of language in terms of applied and theoretical linguistics. The integrative approach that involves corpus linguistics and computational linguistics is necessary for that purpose. Thus, there is a growing need to study and analyze lexical-grammatical associations to identify which lexis, in particular, occurs with which pattern more frequently; this is to serve both grammar and lexicon. In addition, it is important to analyze and study grammatical-grammatical associations to identify the relationships that arise when certain patterns are combined in terms of grammaticality, ungrammaticality, function, and use.

1.2 Purpose of the Study

The main aim of this study was to explore and analyze language structures in terms of two associations: lexical-grammatical associations and grammatical-grammatical associations. Concerning lexical-grammatical associations, the study aimed to explore the differences in frequencies among four English linking verbs (be, seem, look, feel). In other words, it aimed to identify which verb occurred more frequently. Concerning grammatical-grammatical associations, the study aimed to discover differences in frequencies between two relative pronouns used as a subject in essential (restrictive) adjective clauses. In other words, the study attempted to find which pronoun occurred more frequently.

1.3 Questions of the Study

In light of the study's purpose, the following study questions were generated:

1- To what extent do the linking verbs differ in their occurrences in the selected textual structures?
   a. What is the most frequent linking verb occurring in the selected textual structures?
   b. What is the least frequent linking verb occurring in the selected textual structures?

2- To what extent do the subject relative pronouns differ in their occurrences in the selected textual structures?
   a. What is the most frequent subject relative pronoun occurring in the selected textual structures?
   b. What is the least frequent subject relative pronoun occurring in the selected textual structures?

1.4 Delimitations of the Study

This study was delimited to two types of linguistic analysis: lexical-grammatical analysis and grammatical-grammatical analysis. The lexical-grammatical analysis involved four English linking verbs: *be, seem, feel,* and *look*. The linking verbs are supposed to be followed immediately by adjectives. The verbs were investigated in all their possible forms. The second type was grammatical-grammatical associations. This analysis involved two relative pronouns: *who* and *that*. These two relative pronouns were limited to their use for people and as a subject of the dependent adjective clauses in their restrictive (essential) type. The subject of the texts was delimited to health issues and all materials were electronically achieved. The study was also delimited to written form patterns in one restricted topic concerning health-life problems.

2. REVIEW OF LITERATURE

2.1 Introduction

This section deals with the review of literature related to the topic of the current research paper. The review sought to shed light on the concept of corpus linguistics and to provide a brief presentation of its historical background. Then the relationship between the corpus-based approach and the use of computers was discussed in light of the importance of that relationship in

the field of linguistic research. Afterward, the role of concordances and the issues attached to it, such as collocations and KWIC, were investigated. Two main issues were also discussed in this section because they contributed to purposes of the current study; they were lexical-grammatical analysis and grammatical-grammatical analysis. Concerning lexical-grammatical analysis, linking verbs were discussed and elaborated. At the level of grammatical-grammatical analysis, adjective clauses were explained and identified.

2.2 Corpus Linguistics between Concept and History

There cannot be one ideal paradigm through which linguistics to different languages is undertaken as an axiom. Over centuries, endeavors never stopped, nor did they integrate to achieve consensus. Thus, decades of "studies and investigations" always result in newer trends in linguistics, particularly at two main levels: structure and use. Previously, pioneers in the field of language were concerned mainly with studying the surface level of language—in other words, how small units combine to form larger units with regard to constituent class. Then these associations are judged in terms of grammaticality and ungrammaticality.

Recently, there has been a strong shift toward a new trend in language studies, (Biber, et al., 1998). The new approach to linguistics is to use corpora to analyze associations at various levels: lexical-lexical, lexical-grammatical, grammatical-grammatical, and linguistic-non linguistic. Particularly worthy of note, concepts such as function and use were focused on rather than structures and rules. More importantly, brand-new areas have been investigated as factors affecting language use, such as register, genre, and dialects. For example, the language of fiction is compared to the language of non-fiction to explore the distinctive characteristics of both uses.

Traditionally, a large number of investigations were dedicated to language analyses. The analyses were usually done manually and were devoted to issues of the classical approach to language study, as mentioned above. Of course, the compilation of corpora was neither qualitative nor quantitative enough, as there were no computerized devices. Moreover, the labor of analysis itself was greatly time-consuming and the results were relatively inaccurate. In other words, the use of corpora is not a new method for language studies, but was limited and incomprehensive with regard to accuracy, variety, language function, qualitative vs. qualitative nature, and generalizations.

This leads us to consider what is meant by corpus linguistics. McEnery, et al. (2006) and Samson (1995:4) define corpus as "a body of naturally occurring language", So it seems, as asserted by Wisniewski (2007:1), that corpora refers to a language in its two forms—written and spoken—that involve millions of words. Chen (2004) supports the same idea and adds that corpora are authentic and electronic language databases that can be downloaded from the Internet or set on desktops as software. Gries (2009) postulates that corpus linguistics is a controversial notion and can be thought of as a method, model, and/or theory. Moreover, a corpus is thought of as a compilation of more than one text used for empirical linguistics (Mcenery & Wilson, 1996). On the other hand, Meyer (2002) claims that corpora are texts or part of texts; consequently, they are used in relation to the functional approach to language analysis.

2.3 Computer and Corpus-Based Approach in Linguistic Research

The era of technology evolution yielded a pivotal revolution in the field of computer sciences. It was to be expected that this development would affect all fields, especially linguistic research. Therefore, specialists described a very important onset years ago through which the use of computers became essential in analytical investigations. This resulted in what we today call the corpus-based approach to linguistic research. In the present time, as maintained by McEnery, et.al, (2006) and Samson (1995), the corpus-based approach implicitly involves the use of computers widely and effectively because of the ample characteristics that computer programs have. Thus, as a result, an underlying integration occurred between the corpus-based approach and computational linguistics to provide more respective results of analyses with regard to many factors, such as representativeness, cost, accuracy, and generalizations.

Thus, the corpus-based approach and computational linguistics NLP are broadly dedicated to studying the characteristics of grammatical rules, language function, pragmatics, language teaching (Chang & Chang, 2004), language learning and acquisition, and discourse analyses, and are also devoted to lexicography development (Mcenery & Wilson, 1996). Greenbaum (1996) investigated the impact of computer use on corpus linguistics with regard to students' learning, specifically, the use of phrasal lexical. He maintained that software is very effective in enabling the analyst to determine the associations in a learner's writing. He also asserted that the use of computers in relation to corpus linguistics provided rich statistical measures regarding quantitative characteristics.

Another framework, which was presented by Moon (1998), ensured the change in lexicology and lexicography by using a large number of corpora and corpus linguistics. The most important piece of information that can be obtained from the corpus linguistics approach is word frequency with regard to lexical associations. This leads us to the importance of concordances that are used to derive collocations. On the other hand, collocations—word associations in context—are an essential method of counting the commonsense occurrences of a word across texts. In this regard, Moon (1998) claimed that the collocation-based approach of corpora is a promising device used to detect lexical behaviors.

In sum, the integrative approach that involves computational linguistics and corpus linguistics is regarded as an ideal axiom of the analytical model to language investigations. In this regard, this approach suits the two main branches of linguistics, i.e., theoretical linguistics and applied linguistics. Consequently, we can say that the use of computers is greatly sufficient in the field of lexical-grammatical associations and grammatical-grammatical associations, which is the main aim of our current research.

2.4 Concordances

The previous discussion stated a claim that corpus linguistics plays a pivotal role in language investigations with the help of NLP. Also, we stated that corpus or corpora comprise a body of language and collections of a huge number of spoken and written patterns that are compiled and classified according to certain regulations and not haphazardly. In other words, today, we can find types of corpora such as learner corpora, diachronic corpora, general corpora, written corpora, etc. (Greenbaum, 1996).

Now, corpus or corpora provide us with unlimited numbers of concordances. Concordances are defined as lists of words in a context that means that a word appears in combination with other words in lines. This combination is called collocation and is defined by Wouden (1997:5) as follows: "COLLOCATION is a term that refers to the mechanism, or fact, that certain words are regularly found in the company of other words", For example, certain verbs are used with certain nouns; we say "make a decision" not "take a decision". Consequently, KWIC—a key word in a context—is an important device used to benefit from collocations to analyze the characteristics of different language associations (Renouf & Kehoe, 2006).

2.5 Areas of the Use of Concordances

2.5.1 Analysis of Lexical-Grammatical Associations

The relationship between lexis and grammar is worth being studied. In this respect, Schonefeld (2001) stated that although lexicon and grammar are two aspects of a language, and some emphasize lexicon while others emphasize grammar, the interface between them remains a plausible notion.

However, as mentioned above, the corpus-based approach in relation to computational linguistics provides a paradigm through which to investigate many aspects of language associations. More specifically, specialists in this field may make use of certain software to analyze lexical-grammatical associations by the means of collocations and KWICs. This refers to the exploration of the relationship between grammar and lexis. In other words, this type of analysis helps in identifying the occurrences of certain grammar with certain vocabulary. For example, Biber, et al.

(1998) argue that two words (little, small) may appear differently with respect to collocations even though they have almost the same meaning. This means that the two words have distinctive grammatical features. The former is used with mass nouns, whereas the latter is used with count nouns. Consequently, the use of a key word within collocations is a device that helps the analyst identify grammatical properties and the relationship between lexis and grammar.

2.5.1.1 Linking Verbs

Talking about verbs means to deal with word classes such as nouns, verbs, adjectives, etc. Verbs are also called lexical words or contentives. They refer to "doing" because they stand for the process in meaning (Freeborn, 1994). Verbs are classified into various categories: tensed (finite), non-tensed (non-finite), regular, irregular, transitive, and intransitive. Often, verbs—if transitive—are followed immediately by NP object, while if they are intransitive, they are followed by adjuncts, i.e., PP or AP. The English language has what is called a linking verb, or copulative verb, which, as stated by Leech (2006), is a predicate that is immediately followed by an adjective and that might appear in another fashion when followed by any complement.

The English language has 12 linking verbs: *be, feel, look, smell, sound, taste, appear, seem, become, get, turn,* and *grow.* For instance, one may say "The weather became warmer", "Do you still feel hungry?" and "The future is looking good". It is worth noting that the verbs "look" and "appear" mean "seem" in this case and that the verbs "get", "turn", and "grow" mean "become" (Azar, 2002).

2.5.2 Analysis of Grammatical-Grammatical Associations

Grammatical-grammatical association is one of the richest areas through which corpus linguistics can present generalized patterns in relation to different language uses. In addition, this type of analysis sheds light on the principles that govern structures in combination. The contribution of corpus linguistics is clear regarding grammaticality and ungrammaticality counted for native speakers' production of language, whether in its written or spoken form. In this regard, McEnery, et al. (2006) stated that corpora have a strong influence on recently published references of grammar. He also maintained that corpus-based grammar can explore the distinctive features that exist between spoken patterns and written patterns using registers.

The combination of corpus linguistics and computational linguistics as stated by Sampson (1995) deals with a rich analysis to identify properties of surface grammar, i.e., word classes and patterns, and the features of logical grammar in relation to the semantic aspect. An important issue related to this aspect was discussed by Renouf (2006), who established a correlation between the use of corpus linguistics and grammatical issues in terms of the characteristics of a reference grammar and corpus linguistics methodology relevant to writing grammar and corpus linguistics research. Several features can be achieved through an analysis of corpora, whether written forms or spoken forms. For example, this type of analysis may help in identifying patterns, such as dependent clauses, within structures—in other words, what phrases precede and what phrases follow the dependent clause in a construction.

2.5.2.1 Adjective Clauses

As defined by Todd (1987:63), a clause "is a group of words which contains a finite verb", Clauses in general are of several types; they are classified mainly into *independent* clauses, which are sequences of words that can stand by themselves as separate sentences, and *dependent* clauses, which are considered a part of a sentence in the sense that they can't stand in isolation. Dependent clauses are of three types: *adjective clauses, noun clauses,* and *adverb clauses* (Azar, 2002).

Adjective clauses usually modify and describe a noun that is mentioned in the higher position in a sentence by providing further information; they are also called relative clauses (Traugott & Pratt, 1980). The adjective clause is combined with the main clause, *independent*, by relative pronouns: *who, that, which, whom, whose, where*, and *when.*

However, relative pronouns have various functions; for instance, *who, that,* and *which* can be used as subjects of the adjective clause: *who* is used for people, *which* is used for things, and *that* is used for both people and things. Furthermore, the adjective clause is classified into two additional sub-categories with respect to necessity.

If the adjective clause is unnecessary in the sense that it adds additional information only about the noun it describes, commas are used and the clause can be omitted. It is called *non-essential* or *non-restrictive*; an equivalent term is "parenthetical". Non-essential clauses permit all pronouns except *that*.

In contrast, if the adjective clause is necessary to the noun it identifies, no commas are used, and it cannot be omitted. This type of clause is called *essential* or *restrictive*; an equivalent term is "contrastive". When an essential clause exits, any possible pronoun may be used.

3. METHODOLOGY AND PROCEDURE

Introduction

This research paper aims to explore the lexical-grammatical associations and grammatical-grammatical associations using the corpus-based approach. The objective of this research is to find the frequency of four linking verbs in the English language: *be, seem, look,* and *feel*. It also aims to identify which of those verbs is more common. The second aim of the current research is to discover the frequency of two relative pronouns used in dependent adjective clauses—*who* and *that*—and to determine which pronoun is more frequent.

3.1 Research Design

This study is descriptive and analytic. Throughout this research, lexical-grammatical associations and grammatical-grammatical associations were investigated using the corpus-based approach. Concerning lexical-grammatical associations, three linking verbs were investigated with respect to frequent occurrences in the texts. On the other hand, grammatical-grammatical associations were investigated with regard to two relative pronouns used as a subject in dependent adjective clauses. Frequencies of all target words were determined to explore common occurrences.

3.2 Data Sources

The study data were collected from various sources. Data concerning the review of literature was collected from electronic sources and publications of books and references. The researcher also made use of the course-pack and lectures presented by Professor Mahmoud Saleh. Electronic materials were a helpful source concerning text compilation. Data concerning textual analysis were collected from the Internet using the downloading technique; then all texts were saved and documented by name in folders.

3.3 Study Samples

The samples of the current study involved 17 articles, papers, and investigations. All materials were electronic in the sense that they were obtained from Internet websites and then compiled as software. All samples were of one type, i.e., health problems and lifestyle. They were collected from newspapers and various health institutions' websites. The following table illustrates the samples.

N	Article	N W W	N V P	N T P
1	BBC News	476	209	454
2	Canada Free Press	738	346	741
3	Daily Mirror	260	139	240
4	Health MSN	936	442	927
5	Lancaster Guardian	1138	421	1188
6	Mental Health Anxiety	1755	653	1736
7	Msn Lifestyle	672	293	653
8	National Health Service	419	199	417
9	National Inst. Child	889	371	869

10	New York Times	1175	448	11211342111
11	Problems of Life -Fear	635	266	615
12	Problems of Life-Guilt	1239	410	1228
13	The Associated Press	830	422	809
14	The Guardian	1002	429	994
15	The Independent	1021	378	907
16	The Telegraph	568	223	486
17	Times on Line	6011	1042	2602
	Total	19764	6691	16000

TABLE 1: Samples of the study.

N W W = Number of words in Word Processing Program
NVP = Number of Vocabulary in Plain text
NTP = Number of Tokens in Plain Text

3.4 Instrumentation and Procedures

Purposes of the current research paper is to find the frequency of two types of associations: 1) lexical-grammatical associations, by finding the higher frequency among four linking verbs: *be, look, feel,* and *seem*, and 2) grammatical-grammatical associations. This purpose involved two relative pronouns (who, that) used primarily as a subject in dependent adjective clauses.

The researcher used a software program called SCP (Simple Concordancer Program) to find the frequencies of the research targets.

First Phase

The first phase was the compilation of the data. For purposes of this study, 17 electronic sources were searched to obtain written texts in the field of health. The materials averaged between 150 and 1000 words. The overall number of words of the entirety of the materials was 19,764.

Second Phase

After compilation of the texts, the data were manipulated and converted to Plain Text. The conversion process was carried out to enable the SCP to deal with the texts.

Third Phase

In this phase, the researcher dealt with the materials one by one to count the frequencies of:

1-The linking verb (be) in its seven forms: is, was, are, were, be, been, being.
2-The linking verb (seem) in its four forms: seem, seems, seemed, seeming.
3-The linking verb (look) in its four forms: look, looks, looked, looking.
4-The linking verb (feel) in its four forms: feel, feels, felt, feeling.
5-The relative pronoun (who) as a subject of the adjective clause.
6-The relative pronoun (that) as a subject of the adjective clause.

To achieve the above objectives, the researcher used the option (word) and KWIC available in the program. Then the project statistics of each material were found. This research adopted raw count type to extract the word frequency.

3.5 Data Analysis and Statistics

The current study utilized the following statistical methods to analyze the data:

- **The Simple Concordancer Program**
- **Collocations**
- **KWICS**
- **Descriptive Statistics:**
  1. Frequencies
  2. Averages

4. RESULTS AND DISCUSSION

4.1 Introduction

The current research aimed to identify the two types of associations. The first type was lexical-grammatical associations and the second type was grammatical-grammatical associations. Concerning the first type, the study aimed to explore the frequencies of the four linking verbs (be, seem, look, feel) by counting their total occurrences in the entirety of the texts. The second type was concerned with exploring the frequencies of the two subject pronouns (who and that) used as a subject in restrictive adjective clauses. Thus, this section was divided into two sub-sections. The first sub-section dealt with lexical-grammatical associations, and the second sub-section dealt with grammatical-grammatical associations.

4.2 Section (A) Frequency of Linking verbs

Article	N W W	N V P	N T P	N (be)
BBC News	476	209	454	9
Canada Free Press	738	346	741	9
Daily Mirror	260	139	240	2
Health MSN	936	442	927	10
Lancaster Guardian	1138	421	1188	18
Mental Health Anxiety	1755	653	1736	29
Msn Lifestyle	672	293	653	6
National Health Service	419	199	417	10
National Inst. Child	889	371	869	11
New York Times	1175	448	11211342111	15
Problems of Life -Fear	635	266	615	12
Problems of Life-Guilt	1239	410	1228	28
The Associated Press	830	422	809	5
The Guardian	1002	429	994	13
The Independent	1021	378	907	12
The Telegraph	568	223	486	2
Times on Line	6011	1042	2602	28
Total	19764	6691	16000	219

TABLE 2: Frequency of Verb (Be).

The results in table (2) indicate the distribution of verb (be) among the whole of articles. The table shows that the overall number of verb (be) was 219. More specifically, verb (be) represents 3.27 % of the total number of NVP.

Article	N W W	N V P	N T P	N (seem)
BBC News	476	209	454	0
Canada Free Press	738	346	741	0
Daily Mirror	260	139	240	0
Health MSN	936	442	927	1
Lancaster Guardian	1138	421	1188	1
Mental Health Anxiety	1755	653	1736	1
Msn Lifestyle	672	293	653	0
National Health Service	419	199	417	0
National Inst. Child	889	371	869	0
New York Times	1175	448	1134	0
Problems of Life -Fear	635	266	615	3
Problems of Life-Guilt	1239	410	1228	0
The Associated Press	830	422	809	0
The Guardian	1002	429	994	0
The Independent	1021	378	907	0
The Telegraph	568	223	486	0
Times on Line	6011	1042	2602	3
Total	19764	6691	16000	9

TABLE 3: Frequency of Verb (Seem).

As indicated in table (3), the linking verb (seem) represents 0.13 % of the overall number of the NVP. The distribution of the verb among the texts achieved (9) as a total number of the frequency of the verb.

Article	N W W	N V P	N T P	N (Look)
BBC News	476	209	454	0
Canada Free Press	738	346	741	0
Daily Mirror	260	139	240	0
Health MSN	936	442	927	0
Lancaster Guardian	1138	421	1188	0
Mental Health Anxiety	1755	653	1736	0
Msn Lifestyle	672	293	653	0
National Health Service	419	199	417	0
National Inst. Child	889	371	869	0
New York Times	1175	448	1134	0
Problems of Life -Fear	635	266	615	0
Problems of Life-Guilt	1239	410	1228	0
The Associated Press	830	422	809	0
The Guardian	1002	429	994	1
The Independent	1021	378	907	0
The Telegraph	568	223	486	0
Times on Line	6011	1042	2602	1
Total	19764	6691	16000	2

TABLE 4: Frequency of Verb (Look).

According to table (4), the frequency of the verb (look) in all texts was only (2). This indicates that the distribution of the verb among the whole texts is 0.02% with regard to NVP.

Article	N W W	N V P	N T P	N (Feel)
BBC News	476	209	454	0
Canada Free Press	738	346	741	0
Daily Mirror	260	139	240	0
Health MSN	936	442	927	0
Lancaster Guardian	1138	421	1188	0
Mental Health Anxiety	1755	653	1736	5
Msn Lifestyle	672	293	653	0
National Health Service	419	199	417	0
National Inst. Child	889	371	869	0
New York Times	1175	448	1134	0
Problems of Life -Fear	635	266	615	2
Problems of Life-Guilt	1239	410	1228	6
The Associated Press	830	422	809	0
The Guardian	1002	429	994	2
The Independent	1021	378	907	0
The Telegraph	568	223	486	0
Times on Line	6011	1042	2602	1
Total	19764	6691	16000	16

TABLE 5: Frequency of Verb (Feel).

The results in table (5) indicate the distribution of the verb (feel) among the whole of articles. The table shows that the overall number of verb (feel) was 16. More specifically, the verb (feel) represents 0.23 % of the total number of NVP.

4.3 Section (B) Frequency of Relative Pronouns

Article	N W W	N V P	N T P	N (who)
BBC News	476	209	454	8
Canada Free Press	738	346	741	3
Daily Mirror	260	139	240	0
Health MSN	936	442	927	3
Lancaster Guardian	1138	421	1188	3
Mental Health Anxiety	1755	653	1736	2
Msn Lifestyle	672	293	653	0
National Health Service	419	199	417	1
National Inst. Child	889	371	869	14
New York Times	1175	448	11211342111	6
Problems of Life -Fear	635	266	615	1
Problems of Life-Guilt	1239	410	1228	5
The Associated Press	830	422	809	3
The Guardian	1002	429	994	4
The Independent	1021	378	907	5
The Telegraph	568	223	486	5
Times on Line	6011	1042	2602	7
Total	19764	6691	16000	70

TABLE 6: Frequency of Pronoun (who).

Data in table (6) illustrate that the total number of the relative pronoun (who) is 70.

Article	N W W	N V P	N T P	N (that)
BBC News	476	209	454	0
Canada Free Press	738	346	741	0
Daily Mirror	260	139	240	0
Health MSN	936	442	927	2
Lancaster Guardian	1138	421	1188	1
Mental Health Anxiety	1755	653	1736	11
Msn Lifestyle	672	293	653	0
National Health Service	419	199	417	1
National Inst. Child	889	371	869	1
New York Times	1175	448	11211342111	5
Problems of Life -Fear	635	266	615	3
Problems of Life-Guilt	1239	410	1228	1
The Associated Press	830	422	809	0
The Guardian	1002	429	994	1
The Independent	1021	378	907	0
The Telegraph	568	223	486	0
Times on Line	6011	1042	2602	0
Total	19764	6691	16000	26

TABLE 7: Frequency of Pronoun (that).

The results in the above table indicate that the pronoun (that) is used 26 times in total.

5. FINDINGS OF THE STUDY

The results will be discussed according to the two types of associations investigated throughout this research paper.

5.2 Findings Concerning Lexical-Grammatical Associations

The tables displayed in this section reveal that the whole number of the selected texts was 6691 via types in the plain text. The frequency of the verb (be) as a linking verb indicated that it represented 3.27% of the total average. The results for the other linking verbs were as follows:

seem represented 0.13%, *feel* represented 0.23%, and *look* represented 0.02%. These results relatively indicated that the use of (be) as a linking verb across lexical-grammatical associations was more frequent as compared to the other linking verbs. More specifically, the difference among all verbs is not of a high value; in other words, the average of all verbs, especially (look, seem, and feel), is convergent. However, the verb (look) was the least frequent among all verbs.

5.3 Findings Concerning Grammatical-Grammatical Associations

The results showed that the frequency of the subject pronoun (who) was 70, whereas the frequency of the subject pronoun (that) was 26. The results indicated a significant difference between the uses of the two subject pronouns in relative clauses. In other words, the findings showed that the pronoun (who) occurred more frequently than the pronoun (that) in essential (restrictive) adjective clauses.

5.4 Summary and Conclusions

The current research paper investigated two areas of linguistic analysis. There were two main purposes for developing this paper. The first purpose was related to the analysis of lexical-grammatical associations. In this respect, the frequencies of four linking verbs were identified to explore differences among these verbs with respect to their occurrences in the texts. The verbs selected for this research were *be, seem, look,* and *feel*. All forms of the verbs were investigated for an exploration of their frequency.

The second purpose of this study was to analyze grammatical-grammatical associations. Thus, two relative pronouns used to introduce adjective clauses were investigated. In this respect, the frequencies of the selected pronouns were identified to explore differences among them. The purpose was to find which pronoun occurred more frequently than the other one. The adjective clauses were delimited to the essential (restrictive type), and the pronouns were delimited to their use as a subject of the relative clause.

Two approaches were used for purposes of this research: the corpus-based approach and the computational linguistics approach. Thereby, the Simple Concordancer Program was used to identify the differences in frequencies using collocations and KWICS. Two essential research questions were developed for purposes of this study. 1- To what extent do the linking verbs differ in their occurrences in the selected textual structures? 2- To what extent do the subject relative pronouns differ in their occurrences in the selected textual structures?

The findings concerning lexical-grammatical associations indicated that there were no significant differences between the occurrences of the linking verbs. However, the verb (be) occurred relatively more frequently than (*seem, look*, and *feel*). On the other hand, the verb (look) was the least frequent verb among all verbs. The findings concerning grammatical-grammatical associations indicated significant differences between the occurrences of the subject relative pronouns. More specifically, the pronoun (who) occurred more frequently than did the pronoun (that).

The current research could shed light on the linguistic structures as a writing style used by different newspapers with respect to two linguistic aspects: lexical-grammatical and grammatical-grammatical associations. Nevertheless, it would be very beneficial to approach some critical comparison between the results of this research and other research findings in the same field. However, there was no research dedicated to exploring the frequencies of the use of the linking verbs and the relative pronouns, at least to my best knowledge. In this matter, the researcher tried a lot of journals to find articles and/or research investigated corpus-based analysis for the purpose of lexical-grammatical associations and grammatical-grammatical associations, of any sort. Thus, it is strongly recommended that this research is duplicated to seek other issues. For example, a research may be conducted to explore the reasons behind the more frequent use of (who) than the use of (that). This research used newspapers as a source of data analysis. It might be useful to duplicate this research taking data from spoken discourse to get insight if there are differences between written and spoken forms.

6. REFERENCES

[1] Azar, B. Understanding and Using English Grammar. New York: Longman, (2002).

[2] Biber, D., Conrad, S. & Reppen, R. Corpus Linguistics: Investigating Language Structure and Use. Cambridge: CUP, (1998).

[3] Chang, J & Chang, Y. (2010, January 21) Computer Assisted Language Learning Based on Corpora and Natural Language Processing: The Experience of Project CANDLE. IWLeL 2004: An Interactive Workshop on Language e-Learning 15 – 23.

[4] Available: http://dspace.wul.waseda.ac.jp/dspace/bitstream/2065/1391/1/02.pdf. [2004].

[5] Chen, Y. (2004). The Use of Corpora in the Vocabulary Classroom. The Internet TESL Journal: Vol. X, No.9. Available: http://iteslj.org/Techniques/Chen-Corpora.html [January 20, 2010].

[6] Freeborn, D. A Course Book in English Grammar. London: The Macmillan Press, (1987).

[7] Greenbaum, S. The international Corpus of English. Oxford: Clarendon Press, (1996).

[8] Gries, S. (2009). What is Corpus Linguistics. Language and Linguistics Compass: Vol.3, pp. 1-17.
Available:http://www.linguistics.ucsb.edu/faculty/stgries/research/CorpLing_Compass.pdf [January 20, 2010]

[9] Leech, G. A Glossary of English Grammar. Edinburgh: Edinburgh University Press, (2006).

[10] Mcenery, T. & Wilson, A. Corpus Linguistics. Edinburgh: Edinburgh University Press, (1996).

[11] McEnery, T., Xiao, R., & Tono, Y. Corpus-Based Language Studies. New York: Routledge, (2006).

[12] Meyer, CH. English Corpus Linguistics: An Introduction. England: Cambridge University Press, (2002).

[13] Moon, R. Fixed Expressions and Idioms in English: A Corpus-Based Approach. Oxford: Clarendon Press, (1998).

[14] Renouf, A. & Kehoe, A. The Changing Face of Corpus Linguistics. Amsterdam: Rodopi, (2006).

[15] Samson, G. English for the Computer. Oxford: Clarendon Press, (1995).

[16] Schonefeld, D. Where Lexicon and Syntax Meet. New York. Mouton de Gruyter, (2001).

[17] Todd, L. An Introduction to Linguistics. England: Longman York Press, (1987).

[18] Traugott, E. & Pratt, M. Linguistics for Students of Literature. Florida: Harcourt Brace Jovanovich, Inc., (1980).

[19] Wisniewski, K. (2007). Corpus Linguistics. Encyclopedia of Language and Linguistics. Oxford: Elsevier. Available: From the World Wide Web: http://www.tlumaczenia-angielski.info/linguistics/corpus-linguistics.htm [January 20, 2010].

[20] Wouden, T. Negative Contexts: Collocation, Polarity and Multiple Negation. London: Routledge, (1997).

7. APPENDICES

Some Axioms Referring to the Results of the Concordancer Processes

Figure 1

that
10 life, previous work has shown that 95% of victims are not
17 bruises, or burns/•/Injuries that are in the shape of an
55 ./Researchers have found that boys and girls living
56 do so again. It is important that preventive measures are
68 the abuse, reassure him/her that you believe him/her, that
69 her that you believe him/her, that he/she is right to tell
70 she is right to tell you, and that he/she is not bad./•/Tell
77 on Hospital Care recommends that children who are
77 as adults. Research indicates that children without
78 suspected of abuse indicates that training the interviewer
78 correct the interviewer and that it is acceptable for the

Figure 2

who
1 As Bones Age, Who's at Risk for Fracture?/By
4 age, this question arises: Who should be treated with
5 especially older white women, who are at highest risk of one
6 bill. Among elderly people who fracture a hip, 10 percent
12 as they were at age 30 but who have not yet become
12 lead to osteoporosis. Who among them would most
24 bone mineral density range who have the highest risk of
26 for a 60-year-old white woman who is 5 feet and 110 pounds,
27 110-pound woman had a parent who suffered a hip fracture,
32 of vitamin D daily. For those who take omega-3 fatty acids,

Figure 3

that
14 future - all the bad things that may happen to me tomorrow
16 evil spirits./* being alone - that no one cares or
24 someone to marry./* death - that old enemy./How to handle
26 the bottom of our minds, so that perhaps we don't even
29 , "Today is the tomorrow that you spent your time
33 actually need is a new mind, that thinks differently! But
38 . Particularly, do things that will help other people in

Figure 4

who
19 it is not deserved./* People who survive a bad accident or
22 . Someone said, (do you know who?), 'The truth will set you
23 - that we forgive the person who hurt us./This may be a
35 the past. But we know a man who can! What we are
39 people./It is like the man who bought a car from a garage
46 There is the story of the man who wrote to the tax office./"

Figure 5

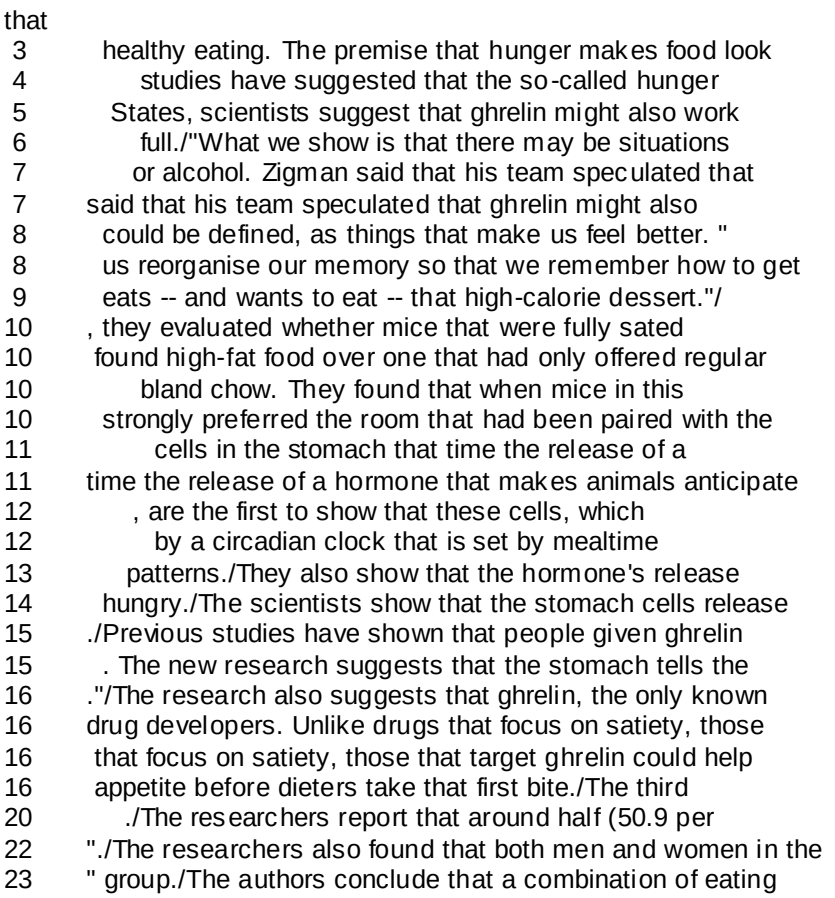

Figure 6

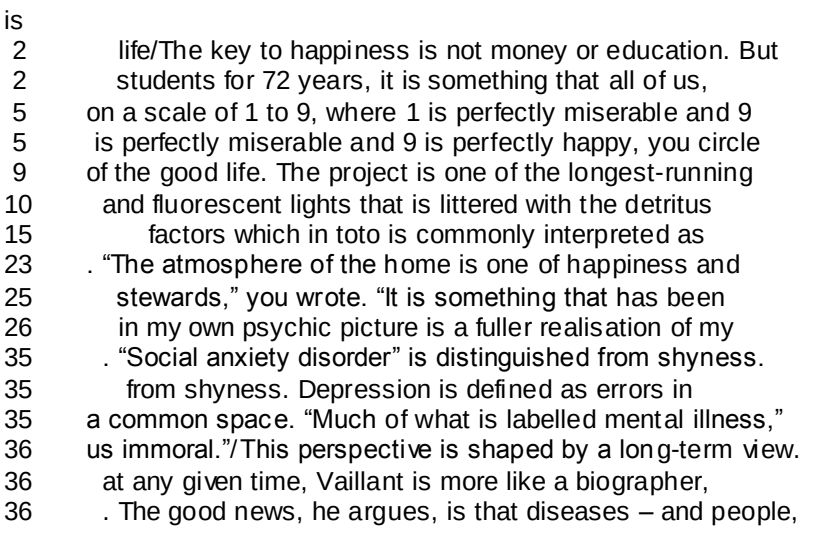

Figure 7

was
11 late eighties – I thought it was a good time to examine
11 to examine both. To do so, I was granted unprecedented
12 told me recently, “but it was contemporary, real-time
12 many decades,” he said, “it was like looking through the
18 presidential Cabinet, and one was President. There was a
18 and one was President. There was a bestselling novelist (
23 But you recovered fully, that was your way. “I could
25 Air Force – “The whole thing was like a game,” you said –
26 on not having any. This was probably because they were
30 started the project when he was 33. He would spend the

Figure 8

seem
33 the person employing them, seem crazy to anyone else. One

Figure 9

are
10 . The files holding the data are as thick as unabridged
10 1993. All that’s missing are the IBM punch cards used
11 the roughly half still living are in their late eighties – I
25 profession. “Our lives are like the talents in the
31 (or “defence mechanisms”) are unconscious thoughts and
33 ./At the bottom of the pile are the unhealthiest, or “
33 to anyone else. One level up are the “immature” adaptations
34 intimacy./“Neurotic” defences are common in “normal” people.
35 . If we use defences well, we are deemed mentally healthy,
36 the “psychotic” adaptations are common in toddlers, and
36 the “immature” adaptations are essential in later
37 with reality that his peers are both moving towards and

Figure 10

were
18 amid the shimmering successes were darker hues. As early as
18 Bock didn’t get it. “They were normal when I picked them
23 , you played all the sports, were good to your sisters, and
24 and ideals.” Basically, they were in a swoon. They
26 and drinking. In 1951 – you were 31 – you wrote, “I think
26 was probably because they were too deeply buried and I
36 , the Grant Study men were twice as likely to use
36 ones, but in middle life they were four times as likely to
36 into old age. When they were between 50 and 75,