

## A Corpus Based Study of Adjectives in Literary and Technical Texts

**Golnar Mazdayasna**

*Yazd University*

*gmazdayasna@hotmail.com*

**Marziyeh Firouzi**

*Yazd University*

*Marziyeh.firouzi@yahoo.com*

### Abstract

In this study, corpus-based techniques were utilized in order to investigate the quantitative and qualitative features of adjectives across literary and technical texts. The corpus was composed of five literary novels and five engineering academic books in English. Seventy paragraphs were randomly drawn from each corpus and the frequency distribution of adjectives with respect to their position (attributive, predicative) and syntactic functions (descriptive, verbal, numeral, etc.) were tallied and summed.

The results revealed that there is a significant difference in the frequency use of adjectives across the two corpora. From a register perspective, the high frequently use of adjectives in technical texts in comparison to literary texts (67.3% and 32.7%, respectively) may be due to the fact that almost all technical texts employ “expository” linguistic features which have a generally “informational purpose”, while most novels employ “narrative” linguistic features which have a direct functional association with the communicative purpose of telling a story of events which have occurred in the past. From a genre perspective, academic texts in engineering tend to be highly informational, non-narrative, and characterized by an impersonal style, whereas novels generally share the same primary communicative purpose of narrating a story whose purpose is to entertain.

**Key words:** Corpus analysis, adjectives, literary texts, technical texts, expository, narrative.

## Introduction

The emergence of corpus analysis has been regarded as a new method of text analysis, aimed at improving language teaching. Linguists use corpora to answer questions and solve problems (Biber, Conrad & Reppen, 1998; Biber & Conrad, 2009; Gavioli, 2005; Meyer, 2004; Reppen & Simpson, 2002). Some of the most revealing insights on language and language use have come from a blend of manual and computer analysis (Gavoili, 2005; Kennedy, 1998). In the field of English for Specific Purposes (ESP) and English for Academic Purposes (EAP), researchers and practitioners seek to understand the linguistic characteristics of specialized registers in English (Biber et al., 1998; Biber & Conrad, 2009; Gavioli, 2005; Meyer, 2004). Corpus analysis techniques have provided evidence about recurring language patterns and about the lexical, grammatical, and lexico-grammatical aspects of language use. One major goal of such research is to design the best possible materials and activities to help students comprehend and produce these registers appropriately (Biber et al., 1998).

As in many countries, teaching EAP has a marginal status in Iranian tertiary education with no clear guidance from institutions or from the Ministry of Science, Research, and Technology as to the expected linguistic or communicative standards (Atai, 2000, 2002; Mazdayasna & Tahririan, 2008; Mazdayasna & Fazilatfar, 2010; Tahririan, 1990). The main aim of EAP classes is to improve technical vocabulary, reading, and translation skills and the textbook seems to represent what the students need to “know” in order to pass their exams. Most English courses make use of a series of unrelated texts (Mazdayasna, 2008).

Most importantly, as it has been echoed in the literature (Biber & Conrad, 2009; Dudley-Evans & St John, 1998; Ewer, 1983), language instructors having a literary background may face some difficulties while teaching ESP courses. Ewer (1983, p. 10) has classified these problems as “attitudinal, conceptual, linguistic, methodological and organizational.” Additionally, ESP students and teachers often have to cope with longer texts than are found in English for General Purpose (EGP) classes (Ewer, 1983; Dudley-Evans & St John, 1998).

In order to overcome these shortcomings, many leading researchers (Biber et al., 1998; Biber & Conrad, 2009; Flowerdew, 2010; Fraser, 2001; Gavoili, 2005; Meyer, 2004; Reppen, 2010) have recommended

the use of specialized corpora which can be effective or fruitful to provide information concerning academic and professional language. Corpus analysis, according to Reppen (2010, p. 67), can be considered as “a valuable resource and a technique for language instructors to investigate a wide range of issues relating to the linguistic characteristics of academic texts; these include vocabulary distributions, the use of collocations and idioms, grammatical characteristics, syntactic complexity, informational density, and rhetorical organizational patterns.”

One of the major contributions of corpus linguistics is in the selection of those features which seem worth teaching in a given pedagogic context (Biber et al., 1998; Biber & Conrad, 2009; Connor & Upton, 2010; Gavoli, 2005; Kennedy, 1998; Reppen, 2010). Adjectives can be considered as one of those lexical features which may vary across different text types (Cao & Fang, 2009; Soler, 2002). Most importantly, there may be qualitative and quantitative differences in different languages and text types. The position of adjectives differs in English and Persian. Persian-speaking students may have some difficulties learning the syntactic functions of adjectives. In Persian, adjectives usually occur in predicative position, and in general they occupy a post-modifying position. However, in English the picture is quite different because adjectives usually occur in attributive position and are used as pre-modifiers. Additionally, in English adjectives also occur in predicative position and occupy a post-modifying position.

To date, few studies have been conducted to investigate the characteristics (morphological, syntactic, and semantic functions) of adjectives (Cao & Fang, 2009; Soler, 2002).

### **The Study**

The main aim of conducting this study was to investigate the frequency distribution and syntactic functions of adjectives across literary and technical texts. Corpus-based techniques make it possible to perform new types of investigations and to conduct research providing multiple perspectives on differences and similarities in language use (Biber & Conrad, 2009). The texts included in the corpus of this study will be analyzed from a register perspective as well as genre perspective. The genre perspective emphasizes the conventional features of whole texts, while register perspective emphasizes variation in the use of linguistic

features (Biber & Conrad, 2009). A crucial part of the corpus-based approach is going beyond the quantitative patterns to propose functional interpretations explaining why the patterns exist (Biber et al., 1998; Biber & Conrad, 2009). Efforts were made to find answers to the following question:

Is there any significant difference between literary and technical texts in terms of using adjectives quantitatively and qualitatively?

The corresponding null hypothesis is that there is no significant difference between literary and technical texts in terms of using adjectives quantitatively and qualitatively.

## **Methodology**

### **The Corpus**

The instrument used in this study to examine the quantitative and qualitative differences across technical and literary texts in terms of using adjectives was corpus analysis techniques. The data were collected from two corpora, namely novels and engineering academic books in English. Each corpus consisted of five literary novels and five engineering academic books. The five engineering academic books were taken from the library of Yazd University and the five literary novels were downloaded from the web site called 'gigapedia'.

While selecting the texts and novels, three criteria were taken into account. First, the latest published engineering academic books and novels were selected; second, the engineering academic books which contained fewer tables, graphs and charts, were selected, and third, the textbooks and novels selected as materials were approximately of the same length. For convenience, the corpus consisting of novels is called the literary corpus and the corpus consisting of engineering texts is called the technical corpus. The materials selected as technical and literary corpora are mentioned below:

#### 1. Literary corpus:

Fette, B. (2004). *Children of Ruin*.

Comish, D. (2004). *A Long Way to Freedom*.

Christian, M. (2006). *Fire and Water*.

Salo, G. (2008). *The Crimson Wrath*.

Reeves-McMillan, M. (2009). *City of Masks*.

2. Technical corpus:

Shooman, M. L. (1984). *Software Engineering: Design, Reliability and Management*.

Ambekar, A. G. (2008). *Mechanical Vibrations and Noise Engineering*.

Stoessel, F. (2008). *Thermal Safety of Chemical Processes: Risk Assessment and Process Design*.

Dunn, C. E. (2007). *Biogeochemistry in Mineral Exploration*.

Markon, S., Kita, H., & Bartz-Beiestein, T. (2006). *Control of Traffic System in Buildings*.

In the next stage, fourteen paragraphs were randomly selected from each text. In this manner, a total number of seventy paragraphs were randomly extracted from five engineering academic books as well as seventy paragraphs were randomly drawn from five novels. Furthermore, in order to avoid the possible effect of inconsistency in the result of the study, each selected paragraph was of the same range of length (approximately, 155 to 165 words). Table 1 and Table 2 display more detailed information regarding each corpus.

Table 1

*Literary corpus*

	Number of paragraphs	Number of words
Romantic novel	14	2301
Social novel	14	2219
War novel	14	2271
Social novel	14	2236
Adventured novel	14	2285
Total	70	11312

Table 2

*Technical corpus*

	Number of paragraphs	Numbers of words
Civil engineering	14	2212
Electronic engineering	14	2240
Chemistry engineering	14	2192
Industry engineering	14	2290
Mechanical engineering	14	2256
Total	70	11190

## Method of Analysis

After selecting seventy paragraphs from each corpus, the next stage was to determine the frequency distribution and syntactic functions of adjectives in each paragraph separately. In other words, the frequency distribution of adjectives along with their syntactic functions (structural properties) were tallied and summed. The syntactic characteristics of adjectives, that is, the distribution of attributive (pronominal position) and predicative (post nominal) adjectives in conjunction with the descriptive, verbal, numeral, demonstrative and compound adjectives were tallied and analyzed across the two corpora.

The numerical data obtained from the two corpora were coded and entered into a computer database. The Statistical Package for Social Sciences (SPSS) software was utilized for statistical operation needed for data analysis. First, the frequency distribution of adjectives with respect to their position (attributive, predicative) and structural properties (descriptive, verbal, numeral, etc.) in each corpus was analyzed. Second, chi-square tests were conducted in order to determine whether the differences in the occurrences of adjectives across the two corpora were statistically significant ( $p < .05$ ).

## Results

In this study, corpus-based techniques were utilized in order to investigate the quantitative (frequency distribution) and qualitative (syntactic functions) features of adjectives across different types of texts. The corpus was composed of five literary novels and five engineering academic books in English. Table 3 displays the frequency distribution of adjectives with respect to their position (attributive, predicative) and syntactic functions (descriptive, verbal, numeral, etc.) in each corpus.

Table 3

*The comparison of the frequency distribution of adjectives across the two corpora*

Syntactic characteristics of adjectives	Literary (Frequency)	Literary (Percentage)	Technical (Frequency)	Technical (Percentage)	P value
Descriptive attributive	375	30.2	865	69.8	.000
Descriptive predicative	140	44.0	178	56.0	.033
-ing attributive	33	37.1	56	62.9	.015
-ing predicative	5	71.4	2	28.6	.257
-ed attributive	36	30.5	82	69.5	.000
-ed predicative	22	55.0	18	44.0	.527
Ordinal	9	20.0	36	80.0	.000
Cardinal	42	42.9	56	57.1	.157
Indefinite	25	32.5	52	67.5	.002
Demonstrative	18	18.2	81	81.8	.000
Compound	8	15.4	44	84.6	.000
Total number of attributive adjectives	444	30.7	1003	69.32	.000
Total number of predicative adjectives	167	45.8	198	54.2	.105
Total number of adjectives	713	32.7	1470	67.3	.000



In both literary and technical corpora, it was observed that adjectives occurred in attributive (pre-nominal) and in predicative (post-nominal) position. Therefore, these adjectives were tallied and analyzed separately in each position and in each corpus. As Table 3 reveals, descriptive attributive adjectives occurred more frequently in technical corpus (69.8%), and less frequently (30.2%) in the literary corpus. Likewise, -ing (verbal) attributive adjective occurred with a high frequency in technical corpus (62.9%) in comparison to literary corpus (37.1%). Similarly, -ed attributive adjectives occurred more frequently in technical corpus (69.5%) in comparison to literary corpus (30.5%). Correspondingly, it can be noticed that the distribution of ordinal adjectives occurred more frequently in technical texts (80%) in comparison to literary texts (20%). The results also revealed that 57.1 percent of cardinal adjectives occurred in technical corpus, while 42.9 percent of cardinal adjectives occurred in literary corpus.

Correspondingly, the results also indicated that indefinite adjectives occurred more frequently in technical corpus (67.5%) in comparison to literary corpus (32.5%). In addition, a comparison of frequencies of demonstrative adjectives observed in literary versus technical texts displayed in Table 3 reveals that demonstrative adjectives occurred with a high frequency in technical texts (81.8%) in comparison to literary texts (18.2%). Additionally, compound adjectives occurred with a higher frequency in technical texts (84.6%) in comparison to literary texts (15.4%).

Furthermore, the results also revealed that some adjectives occurred in predicative (post-nominal) position in both literary and technical texts. Descriptive predicative adjectives occurred more frequently in technical texts (56%) than literary texts (44%). In contrast, -ing participle adjectives in predicative position occurred more frequently in literary texts (71.4%) in comparison to technical texts (28.6%). Similarly, -ed participle adjectives in predicative position occurred frequently in literary texts (55%) in comparison to technical texts (45%).

Moreover, the chi-square test results indicated that there was a significant difference in the use of adjectives namely, descriptive attributive, descriptive predicative, -ing attributive, -ed attributive, ordinal, indefinite, demonstrative, and compound adjectives across technical and literary texts. However, the findings also revealed that there was no significant difference in the use of cardinal adjectives as

well as –ing participle and -ed participle adjectives in predicative position across technical and literary texts.

### **Discussion**

As mentioned earlier, the current study explores the quantitative and qualitative differences across literary and technical texts in terms of using adjectives. The results displayed in Table 3 reveal that the total number of adjectives occurred with a high frequency in technical texts (67.3%) in comparison to literary texts (32.7%). Therefore, the findings of the current study allow the null hypothesis to be rejected confidently. The findings of this study are in consistent with the results which some scholars have found in their studies. For example, Soler (2002) found that the longer the scientific article on Biochemistry the larger the number of adjectives. In four of the five selected papers which he examined, the highest number of adjectives (80%) was found in the Discussion section of scientific articles.

Correspondingly, Cao and Fang (2009) investigated the variations in adjectives use across different text categories selected from the British National Corpus. The findings of their study revealed that variations of adjectives use can not only be applied to the ranking of texts according to degrees of formality, but more importantly to the categorization of texts according to different domains (humanities and arts, medicine, natural science, politics, law and education, social science and technology and engineering). They concluded that “the variations of adjective use seem to be a quiet reliable indicator to categorize different text categories in a meaningful way” (Cao & Fang, 2009, p. 216).

Furthermore, another crucial distinction was observed in the frequency use of attributive adjectives in comparison to predicative adjectives in engineering texts. The total number of attributive adjectives occurred with a high frequency in technical texts (69.3%) in comparison to literary texts (30.7%). The findings of the study conducted by Soler (2002) also revealed that attributive adjectives occurred with a much higher frequency in comparison to predicative adjectives in all the five scientific articles which he had examined.

However, the results of this study also revealed that the total number of predicative adjectives was not as high in technical texts as compared to literary texts. 54.2% of the total number of predicative adjectives

occurred in technical texts and 45.8% occurred in literary texts. This may have stemmed from the fact that, as Cao and Fang (2009) hold, attributive adjectives shows a sense of objectivity, while predicative adjectives show more subjectivity. Since technical texts tend to be highly informational and are written in an impersonal style, attributive adjectives occurred more frequently in engineering texts in comparison to predicative adjectives.

Similarly, the findings of the studies conducted by Yamazaki (2002, as cited in Cao & Fang, 2009) and Rayson, Wilson, and Leech (2001, as cited in Cao & Fang, 2009) using corpus-based techniques, revealed that adjectives occurred more often in written texts than in spoken ones, and more frequently in informative writing than in imaginative writing. To be more specific, adjectives occurred most often in academic prose, reviews and hobbies, while they were less frequent in fiction. Correspondingly, Rittman (2008, as cited in Cao & Fang, 2009) contend that adjectives and adverbs will vary by genre because of their unique patterns of usage in text. Thus, the findings of the present study are in consistent with the results which some scholars have found in their studies.

Most importantly, the texts included in the corpora of this study, can be analyzed from both register and genre perspectives. Many leading researchers (e.g., Biber et al., 1998; Biber & Conrad, 2009) claim that register analysis can be applied to large corpora of written and spoken texts with a focus on exploring the frequency of lexico-grammatical features. Regarding technical and literary corpora of this study, it can be maintained that each of these registers demands its own specific lexical and grammatical features.

From a register perspective, the high frequent use of adjectives in technical texts in comparison to literary texts may be, due to the fact, that academic texts in engineering have a generally informational purpose, as opposed to novels, whose purpose is to entertain. The content is generally factual, not imaginative. Similarly, most technical texts employ “expository” linguistic features (e.g., frequent nouns and long words, with most nouns, being modified by attributive adjectives or prepositional phrases) (Biber & Conrad, 2009, p. 113). On the other hand, according to Biber and Conrad (2009, p. 155), “novels from all periods are similar in their frequent use of grammatical devices associated with narration, reported speech and direct portrayal of

dialogue (e.g., past tense verbs, perfect aspect, third person pronouns, reporting clauses).

From the genre perspective, a distinction can also be made between literary and technical texts utilized as corpora in this study. Grabe (2002) argues that there are two families of macro-genres, the “narrative” and the “expository”. Correspondingly, Biber and Conrad (2009) and Grabe (2002) promulgate that most novels employ “narrative” linguistic features which have a direct functional association with the communicative purpose of telling a story of events which have occurred in the past. These features are found in both modern novels as well as eighteenth-century novels.

In contrast, expository texts, according to Grabe (2002, p. 253), “are assumed to draw on a frame of logic that is not readily open to a range of interpretations but follows a logic that is displayed by the expository text itself (whether true or not, whether accurate or not).” Almost all expository texts employ “expository” linguistic features (e.g., nouns, relative clauses, attributive adjectives) which have a generally “informational purpose” (Biber & Conrad, 2009, p. 113), as opposed to, novels whose purpose is to entertain. The primary purpose of engineering textbooks is to present information as opposed to the personal purposes of literary texts. Adjectives commonly modify nouns, so they add to the informational density of registers like academic prose. Adjectives are much more common in engineering texts because they are an important device used to add information to noun phrases. Therefore, another justifiable reason for the frequent use of adjectives in engineering textbooks as compared to literary novels may be attributed to the inherent nature of narrative and expository text types. The important point for this distinction is highlighting the specific characteristics of these genres that are associated with particular linguistic features.

### **Conclusion**

The most salient findings of this study are that there is a close relationship between the communicative purpose of each text type and the use of adjectives. Adjectives are much more common in engineering texts because they are an important device used to add information to noun phrases. The primary purpose of engineering text is to present information as opposed to the personal purposes of literary text. Adjectives are used more frequently in expository texts to introduce new

technical information which is unknown and difficult to understand. Textbooks in engineering disciplines will also include procedural information (Biber & Conrad, 2009), while textbooks in the humanities will often include persuasive discussion and some narratives. Likewise, another striking difference is in the prominent use of attributive adjectives as opposed to, predicative adjectives which have the pretension to show a sense of objectivity in transmitting knowledge in engineering texts.

Finally, corpus-based studies, such as the current study often have implications for a variety of issues related to language and education. Generally speaking, corpus based studies can provide valuable insights which are useful for language pedagogy and course design. Based on the findings of this study, material designers should take into consideration the different characteristics of adjectives (morphological, semantic and syntactic) while selecting and designing materials across different academic disciplines. Corpora of specialized texts can be a very useful technique in providing indications about key lexical, grammatical or technical issues to deal with in ESP and EAP classes.

Another related pedagogical application of register and genre descriptions can be used in the writing class. It would be beneficial for language instructors in whatever context to increase their awareness of how registers and genres can differ from one situation to another, helping their students to develop a similar awareness of register differences, specifically an awareness of the linguistic norms expected in academic written registers.

### References

- Ambekar, A. G. (2008). *Mechanical vibrations and noise engineering*. New Delhi: Prentice-Hall of India Private Limited.
- Atai, M. R. (2000). *ESP revisited: A reappraisal study of discipline-based EAP programs in Iran* (Unpublished doctoral dissertation). Isfahan University, Iran.
- Atai, M. R. (2002). Iranian EAP programs in practice: A study of key methodological aspects. *Sheikhbahaee ELT Journal*, 1 (2), 1-15.

- Cao, J., & Fang, C. A. (2009). Investigating variations in adjective use across different text categories. *Research in Computing Science*, 41, 207-216.
- Christian, M. (2006). *Fire and water*. London: Lulu Enterprises.
- Commish, D. (2004). *A long way to freedom*. London: Lulu Enterprises.
- Connor, U., & Upton, T. A. (2010). *Applied corpus linguistics: A multidimensional perspective*. Amsterdam: John Benjamin Publishing Company.
- Dudley – Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Dunn, C. E. (2007). *Biogeochemistry in mineral exploration*. Oxford: Elsevier.
- Ewer, J. R. (1983). Teacher training for EST: Problems and methods. *The English for Specific Purpose Journal*, 2 (1), 9-31.
- Fette, B. (2004). *Children of ruin*. London: Lulu Enterprises.
- Flowerdew, L. (2010). The argument for using English specialized corpora. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 11-33). Amsterdam: John Benjamin Publishing Company.
- Fraser, S. A. (2001). Vocabulary and the teaching of English for specific purposes. *Integrated Studies in Nursing Science*, 3 (1), 32-37.
- Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam: John Benjamin Publishing Company.
- Grabe, W. (2002). Narrative and expository macro-genres. In A. M. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 241-267). Lawrence Erlbaum.
- Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.

- Markon, S., Kita, H., & Bartz-Beiestein, T. (2006). *Control of traffic system in buildings*. London: Springer London Limited.
- Mazdayasna, G. (2008). *Developing a profile of the ESP needs of Iranian students: The case of students of nursing and midwifery*. (Unpublished doctoral dissertation). Isfahan University, Iran.
- Mazdayasna, G., & Fazilatfar, A. M. (2010). The role of native language in teaching English for specific purposes. *Iranian Journal of Applied Linguistics*, 13 (1), 99-124.
- Mazdayasna, G., & Tahririan, M. H. (2008). Developing a profile of the ESP needs of Iranian students: The case of students of nursing and midwifery. *Journal of English for Academic Purposes*, 7 (4), 277-289.
- Meyer, C. (2004). *English corpus linguistics: An introduction*. Cambridge: Cambridge University Press.
- Reeves-McMillan, M. (2008). *City of masks*. London: Lulu Enterprises.
- Reppen, R. (2010). Academic language: An exploration of university classroom and textbook language. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 65-88). Amsterdam: John Benjamin Publishing Company.
- Reppen, R., & Simpson, R. (2002). Corpus linguistics. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 92-114). London: Oxford University Press.
- Salo, G. (2008). *The crimson wrath*. London: Lulu Enterprises.
- Shooman, M. L. (1984). *Software engineering: Design, reliability and management*. Singapore: McGraw-Hill Company.
- Soler, V. (2002). Analyzing adjectives in scientific discourse: An exploratory study with educational applications for Spanish speakers at advance university level. *English for Specific Purposes*, 21 (2), 145-165.
- Stoessel, F. (2008). *Thermal safety of chemical processes: Risk assessment and process design*. Weinheim: Wiley-VCH.

Tahririan, M. H. (1990). A summative evaluation of teaching English in Iranian universities. *Research Bulletin of Isfahan University*, 6 (2), 1-8.