

A CORPUS-BASED ANALYSIS OF THE MOST FREQUENT ADJECTIVES IN ACADEMIC TEXTS

by **Galip Kartal**

Necmettin Erbakan University

Meram Yeni Yol, Meram, 42090, Konya, Turkey

kartalgali @ gmail.com

Abstract

Based on a mega corpus, The Corpus of Contemporary American English (COCA), this study aims to determine the most frequent adjectives used in academic texts and to investigate whether these adjectives differ in frequency and function in social sciences, technology, and medical sciences. It also identifies evaluative adjectives from a list of a hundred most frequently used adjectives. A total of 839 adjectives, which comprises the list of frequently used adjectives in COCA, were searched using a search engine. 334 of the adjectives were found to appear more frequently in the academic sub-corpus than in other sub-corpora (spoken, fiction, magazine, and newspaper). There was only one adjective that was used more frequently in technology and medical sciences than in social sciences. Some adjectives were very dominant in a specific discipline of academic texts. The frequency of evaluative adjectives in most frequently used 100 adjectives was also listed. It is found that almost 40% percent of the adjectives are evaluative. The results of the study were discussed in terms of frequency effects in language learning and writing in the foreign language as providing learners with corpus data may improve language knowledge and the correct use of adjectives.

1. Introduction

Wiebe (2000) argues that corpora have been used to obtain linguistic knowledge in natural language processing. Thus, the linguistic knowledge on adjectives can be gathered from available corpora. The focus is on the evaluative adjectives as the knowledge of the evaluative language may be beneficial for text categorization and summarization (Wiebe, Bruce, Bell, Martin, & Wilson, 2001). Evaluation, in this study, is used as defined by Hunston and Thompson (2000), who see evaluation as a means of expressing the speaker or writer's attitude and feelings toward the language they produce. There are many linguistic features that can make a sentence evaluative; however, adjectives are the most frequently used and important tool for evaluating a sentence (Marza, 2011). In another study on evaluative and

speculative language, Wiebe et al. (2001) found that the type of subjectivity was more evident in adjectives than in modals and adverbs.

This study is motivated by four facts. First, previous corpus-based studies on adjectives were done with relatively small corpora (Marza, 2011; Samson, 2006). In their literature review on frequency effects in second language acquisition, Kartal and Sarigul (2017) concluded that the number of the studies investigating the frequency effects via mega corpora is rare. Therefore, exploring adjectives in a mega corpus such as COCA might be useful. Second, previous research has proved that a corpus-based study on evaluative adjectives may help increase foreign language students' awareness of adjective types and usage tendencies in different registers. Third, providing students with real data (corpus data) may improve language knowledge and the correct use of adjectives. Last, frequency helps to quantify the usefulness of a word.

2. Background to the study

2.1. Frequency and usefulness

Although frequency in the input is not the only predictor of the usefulness of a word, the literature shows that frequency and usefulness are strongly related to each other. There are some criteria to determine the usefulness of a word. These include frequency, range, availability, coverage, learnability, and opportunism (White, 1988). According to Nation and Waring (1997, p. 17), frequency information ensures that “learners get the best return for their vocabulary learning effort.” Thus, frequency seems to be the most appropriate measure to decide on the usefulness of a word.

2.2. Evaluative adjectives

Evaluation is an “elusive concept” (Hunston & Thompson, 2000), which is sometimes called “appraisal” (Martin & White, 2005) or “stance” (Conrad & Biber, 2000; Hyland, 2005). The fluctuation in terminology is a result of an abundance of parameters used to conduct evaluation. According to Hunston and Thompson (2000), evaluation refers to judgments, feelings, or viewpoints about something. They also delineate three functions of evaluation: expressing an opinion, maintaining relationships, and organizing discourse. Expressing an opinion is a way to understand the value system of the speaker. Secondly, evaluation acts as a bridge between writer and reader. This relationship can be used for manipulation, hedging, and politeness. Finally, evaluation acts as a discourse organizer. In other words, evaluation not

only builds relationships and conveys values, but also helps coherence (pp. 6-9). As Hyland (1998) believes, evaluation is important for interpersonal metadiscourse. As metadiscourse improves coherence in a passage (Aidinlou & Vafaei, 2012), the use of evaluation plays a significant role in the effectiveness of a text. Evaluative adjectives are also important in discourse (Samson, 2006).

Previous research about evaluative adjectives has focused on written and spoken academic genres, particularly research articles, textbooks, and spoken lectures (Samson, 2006; Swales & Burke, 2003). Samson (2006), for instance, conducted a small corpus study in economic discourse and found that evaluative adjectives have more than one function at the same time and that they differ across genres and registers. The functions were “interacting with readers by underscoring the crucial points in their texts and to promote the economists’ findings by asserting that theirs is a correct interpretation of the topics” (p. 243). Swales and Burke (2003) found that adjectival evaluation is used more frequently in the spoken register by investigating evaluative adjectives in different academic registers. Stotesbury (2003) investigated 300 articles published in 51 journals, including 100 articles in humanities, social sciences, and natural sciences. He found that there were more evaluative attributes in articles in humanities and social sciences than in natural sciences. In addition, evaluative adjectives in articles in economics were more numerous than in linguistics articles.

So far, adjectives have been categorized according to morphological, functional, syntactic, semantic, and pragmatic criteria. Kerbrat-Orecchioni’s (1980) classification of adjectives, for instance, relies on pragmatic criteria (see Table 1).

Table 1. Classification of adjectives (Kerbrat-Orecchioni 1980)

Objective	Subjective		
	Emotional	Evaluative	
		Non-axiological	Axiological
Single/married	Sad	Cold	Bad
Male / Female	-	-	-

Kerbrat and Orecchioni (1980) define non-axiological evaluative adjectives, which have a gradual nature without any subjective emotional bias. Axiological adjectives, on the other hand, reflect the speaker’s positive or negative judgment.

After analyzing evaluative adjectives in a corpus, Marza (2011) concluded that “some evaluative dimensions are seen to be more central than others in the genre under study and those recurrent, emphatic lexical patterns of an evaluative nature clearly characterize this kind

of discourse.” Hewings (2004) grouped evaluative adjectives into eight categories after completing a corpus-based analysis. The categories are listed below with positive and negative examples:

- a. Interest (*interesting, tedious*)
- b. Suitability (*good, odd*)
- c. Comprehensibility (*clear, confusing*)
- d. Accuracy (*true, wrong*)
- e. Importance (*useful, meaningless*)
- f. Sufficiency (*sufficient, small*)
- g. Praiseworthiness (*impressive, disappointed*)
- h. Perceptiveness (*sophisticated, unaware*)

2.3. Subjectivity and adjectives

The term ‘subjectivity’ is used to express opinions and evaluations (Wiebe, 1994). Evaluation and speculation are two main types of subjectivity (Wiebe et al., 2001). According to Wiebe and her colleagues, evaluation includes emotions, judgments, and opinions. Speculation is uncertainty. News reporting and forums, in which opinions are expressed, are suitable for subjectivity tagging (Wiebe, 2000) and the use of gradable adjectives plays a crucial role while determining subjectivity.

According to Wiebe, (2000) identifying linguistic clues to determine subjectivity requires comprehensively-coded tools for subjectivity tagging. Similarly, Bruce and Wiebe (2000) found a statistically significant correlation between the existence of an adjective and subjectivity in a sentence. Leech (1989) points out that after nouns and verbs, adjectives is the largest word class in English. Hunston and Sinclair (2000) found a positive relationship between evaluation and adjective behavior.

3. The study

3.1. The aims of the research

This study focuses on academic texts in COCA because “academic writing has gradually lost its traditional tag as an objective, faceless and impersonal form of discourse and come to be seen as a persuasive endeavor involving interaction between writers and readers” (Hyland, 2005, p. 174). The research questions addressed in this study are as follows:

1. Which adjectives are used most frequently in the academic sub-corpus of COCA?
2. Are there any differences between social sciences and technology and medical sciences in terms of frequency and functions of evaluative adjectives?
3. How many of the frequent adjectives in academic texts are evaluative?

3.2. The corpus

This study utilized the Contemporary Corpus of American English (COCA), a contemporary and genre-based corpus. The corpus covers the years between 1990 and 2012. COCA was used for this research because it is free to access, and it is a mega corpus which includes over 450 million words. This means that it has very comprehensive and highly representative data. In addition, its contemporariness, representativeness, genres, and size are all outstanding when compared with other corpora available.

COCA includes five main sub-corpora: spoken, fiction, magazine, newspaper, and academic. The academic sub-corpus has about 83 million words, and the data are obtained from 148 academic journals. The academic part includes history, education, geography/social science, law/political science, humanities, philosophy/religion, science/technology, medicine, and miscellaneous.

3.3. Selection of adjectives

The Corpus of Contemporary American English can be searched using its search engine. However, the totality of data for a specific word category cannot be reached from the search engine. So, the first 5,000 most frequent words in the COCA corpus were taken from <http://www.wordfrequency.info>, a website which supplies frequencies of words within many corpora. A free list of the 5,000 most frequent words in COCA was used, and 839 of the words in this list were adjectives. In other words, 17% of the most frequent words in COCA are adjectives (see Figure 1). Then, from this list of 839 adjectives, the ones most frequently used in the academic division were extracted. The new list, which is the focus of this study, included 334 adjectives.

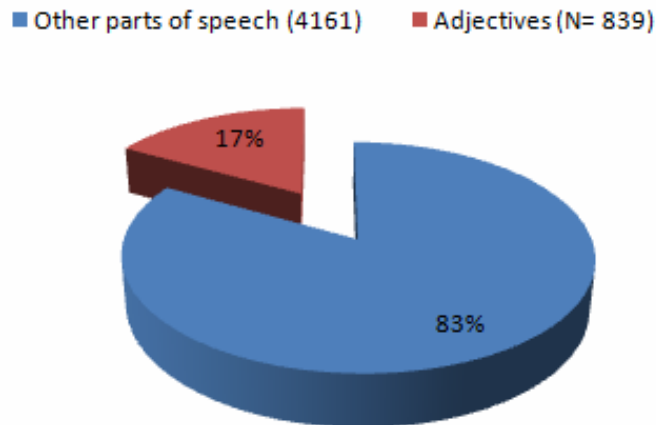


Figure 1. Frequency of adjectives and other parts of speech in the 5,000 most frequent words in COCA

3.4. Results and findings

3.4.1. Research Question One: *Which adjectives are used most frequently in the academic sub-corpus of COCA?*

The results of this corpus-based study revealed that 334 of the 839 adjectives in COCA were more frequently used in the academic sub-corpus when compared to adjectives used in spoken language, fiction, magazines, and newspapers (see Figure 1). In other words, almost 40% of the most frequently used 839 adjectives are mostly found in the academic sub-corpus of COCA. The list of the first one hundred most frequently found adjectives in the COCA academic corpus is provided in Appendix 1.

If we have a look at Hyland (2005) and Bruce and Wiebe (1999), we can conclude that the results of the first research question of this study are not unexpected. Hyland stresses the new feature of the academic text. According to him, academic language is shifting from neutral to a more persuasive way. One of the more important ways of persuasion is using evaluation. In addition, Bruce and Wiebe found a significant relationship between evaluation and adjectives.

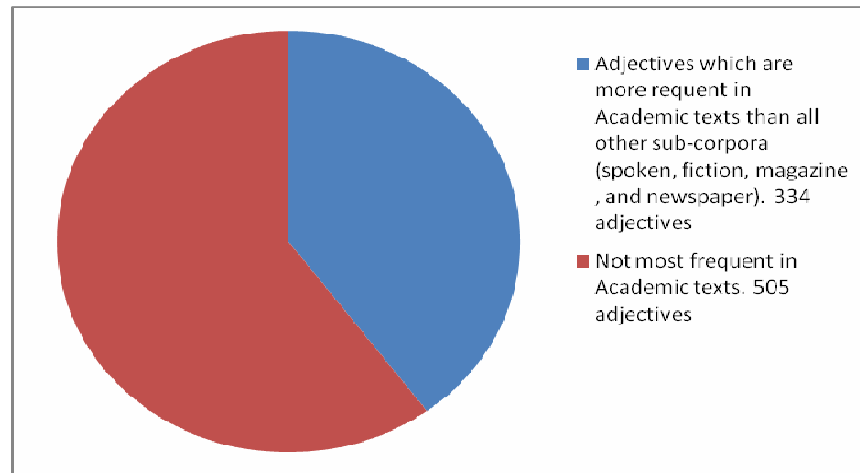


Figure 2: The distribution of 839 frequent adjectives in Academic texts

3.4.2. Research Question Two: *Are there any differences in social sciences and technology & medical sciences in terms of frequency and function of evaluative adjectives?*

All the adjectives which are more frequently found in the academic register were used more in social sciences (history, education, geography/social science, law/political science, humanities, philosophy/religion) than in medicine and technology. The only exception was *environment*, which is used 11,872 times in social sciences and 21,992 times in science/technology and medicine. Appendices 2 and 3 provide the most 100 frequent adjectives in social sciences and science/technology/medicine respectively.

Every fifth adjective (1st, 5th, 10th et al.) within the fifty most frequently found adjectives in academic texts were analyzed with some examples from the corpus. Hence, a total of 11 adjectives were analyzed in depth (see Table 2).

Table 2. The adjectives which were analyzed in depth

Order	Adjectives
1 st	<i>other</i>
5 th	<i>political</i>
10 th	<i>united</i>
15 th	<i>significant</i>
20 th	<i>international</i>
25 th	<i>environmental</i>
30 th	<i>major</i>
35 th	<i>specific</i>
40 th	<i>religious</i>
45 th	<i>low</i>

50th traditional

The most frequent adjective was *other* in academic texts. It was used more frequently in science/technology than any other academic divisions of the corpus. Some examples are:

- 1) ...in a conceptual trough that encourages such yearning for unknown and romanticized greener pastures of **other** times.
- 2) ...in one zone, so that they could really be considered as species of the **other** zone.

It should be noted that, although *other* is used 27,805 times in the science/technology discipline, it is not always used as an adjective. For example;

- 3) ..., in common chimpanzees, males and females have sets of hierarchy independent of each **other**.

The fifth most frequent adjective was *political*. As it can be guessed, the adjective *political* was used in history and law/political sciences much more than in others. It was used only 462 times in medicine, and it was frequently used with the noun *system*.

- 4) Data manipulated by supercomputers are the lifeblood of the modern **political** system.

The next adjective (10th) was *united*. Again, it is used mostly in history and law/political sciences. What makes it very frequently used the word is the fact that it is the official name of America: The *United States of America*. Most of the usage is related to the country's name. *United Nations* is in second place. However, *United* is sometimes used in different ways:

- 5) "The shoeworkers, pianomakers, barbers, hotel and restaurant workers and **United** Textile Workers likewise kept out new immigrants,..."

Significant was the fifteenth most frequently used adjective. When taking a close look at its usage in the corpus, it can be observed that it was mostly used in education and medicine where it is commonly used to report findings of statistical analysis in education. For example;

- 6) # **Significant** differences have been found between males and females on control...
- 7) There was a modest but **significant** increase in the plasma concentrations of vasopressin during upright tilt in patients.

The adjective *international*, with 3,780 usages was the 20th most frequently used an adjective. It was used more in history and law/political sciences than others. The adjective *international* is used mostly before the nouns like monetary system, commercial, relations, standards, etc. The examples are as follows:

- 8) This made the development of a common view of **international** relations even easier.
- 9) ...a synthetic unit, like SDRS, has been proposed as the basis for the **international** monetary system.

The most 25th most frequent adjective was *environmental*. It was the only one that was used more frequently in science/technology and medicine than social sciences (history, education, geography/social science, law/political science, humanities, philosophy/religion). It was used in science and technology over 16,000 times.

10) This reduced pressure on **environmental** resources over large areas.

11) **Environmental** movements cannot prevail until they convince people that clean....

In the current order, the next adjective was *major*. It was mostly used in history and medicine. When the real data is analyzed, it can be easily observed that the adjective is generally related to a research report in the history discipline. Here are some examples from each one:

12) The **major** goals for this first research handbook for social studies were to capitalize

13) ...from or about teachers at all levels, and the dearth of significant and reliable **major** studies conducted on a regional scale in accord with well-known research...

The adjective *specific*, which is used 28,082 times in academic texts, was used more frequently in education and geography/social sciences, for instance:

14) ...elementary and secondary school teachers have a strong liberal arts foundation, as well as specific training in teaching.

15) ...flights a week is the measure of interaction or demand for air service from a **specific** city.

The adjective *religious* is used 25,083 times in the academic sub-corpus. As it can be foreseen, it was used mostly in philology and religious disciplines. *Tradition* is the most commonly used noun with *religious*. Here are some examples from the religious context:

16) Torah as fanaticism and blood-and-conquest, versus modernity (and possibly other **religious** traditions) as peaceable.

17) ... his point is important. It explains the difference between, for example, the **religious** life of a North Asian people and the religious experience of its shamans; ...

The next adjective is *low* with a usage frequency of 23,943. It is used more in geography/social sciences (7,036) and science/technology (4,856) than other disciplines.

18) Eastern Apacheria is a mountainous, arid environment dominated by the Chihuahuan desert at **low** elevations and pine forests at high ones.

19) By maintaining a **low** metabolism and temperature, the cluster of Himalayan honey bees can reduce its food requirements.

The last adjective (50th) is *traditional*. Interestingly enough, this adjective was almost equally distributed across all disciplines of the academic context. Its usage frequency in history, education, geography/social sciences, and humanities were very close to each other.

20) More **traditional** research continues with topics that deal with the teachers' use of questions at various.

21) **Traditional** sit-down restaurants are by-passed in favor of standardized, packaged fare.

22) LeBlanc and Jan McCrary, in a 1983 study, presented two dozen excerpts of **traditional** instrumental jazz to fifth and sixth-grade students...

The analysis of eleven adjectives (every fifth adjective, 1st, 5th, 10th) that are on the list of most frequently used 50 adjectives in the academic texts of COCA revealed that these adjectives have different frequencies and functions in the disciplines of the academic register. This finding is consistent with the results of Stotesbury (2003) to some extent. He found that evaluative attributes in humanities and social sciences are used more frequently than in the natural sciences.

3.4.3. Research Question Three: *How many of the frequent adjectives in academic texts are evaluative?*

The frequency rate of the first 100 adjectives was evaluated and with the evaluative adjectives extracted. There were 35 evaluative adjectives (e.g., *important, significant, difficult*).

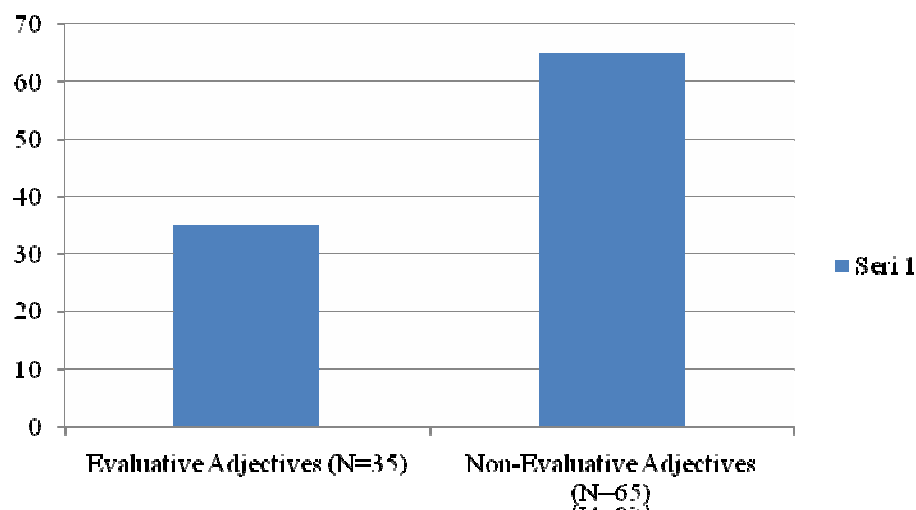


Figure 3: Frequency of evaluative and non-evaluative adjectives in most frequent 100 adjectives

In Marza's (2011) study, who investigated an "untagged corpus of websites owned by both independent hotels and hotels belonging to chains from the United Kingdom and the USA" (p. 105), it was found that 130 of the adjectives in 290 adjective types were purely evaluative from 2000 top frequency types. In other words, 45.2% of all adjectives were evaluative. The present study found very similar results. Out of 839 preliminarily taken adjectives from the top 5,000 frequent words in COCA, 334 of the adjectives were evaluative. The percentage of evaluative adjectives was 39.80%, only 6 percent below the percentage in Marza's study.

4. Discussion and conclusion

This study investigated adjectives with the help of a mega corpus, COCA. The initial research hypothesis was that some adjectives would be more frequent in academic register than in the other sub-corpora of COCA. 17% of the most frequent 5000 words in COCA were adjectives. The corpus analysis revealed that almost 40% percent of the most frequent adjectives were most frequent in academic texts. In addition, the disciplines of the academic register were grouped into two categories: a. social sciences (history, education, geography/social science, law/political science, humanities, philosophy/religion); and science/technology, medicine. There was only one adjective which was more frequent in science/technology, medicine than in social sciences.

It is important for foreign language learners and writers to know which adjectives are frequently used. Thus, EFL learners should be guided to reach authentic use of linguistic items. From this point of view, it can be concluded that providing learners with a list that shows the most frequent adjectives and their functions would be an effective way of helping learners to use those adjectives appropriately.

This study revealed that almost 40% percent of the adjectives in the COCA corpus are evaluative ones. Thus, while using evaluative adjectives in research papers, EFL writers can utilize the appropriate use for any genre and register. Moreover, the use of correct evaluative adjectives is not only important for the genre and the register of the text but also for the correct understanding of the message. In other words, this corpus-based study on evaluative adjectives may increase learners' awareness of adjective types and usage tendencies in different registers.

Previous research has revealed that evaluative adjectives constitute a coherent semantic and syntactic class (Quirk, Greenbaum, Leech, & Svartvik, 1985; Kertz, 2006). There was little focus on the semantic and syntactic analysis of the target adjectives. A further study on these frequent adjectives may focus on semantic and syntactic functions of evaluative adjectives. By this means, the language learners may profit better.

Note

A former version of this paper was presented at the International Conference on the Changing World and Social Research I, held in Vienna-Austria on August 25-28, 2015.

References

- Aidinlou, N. A., & Vafaei, A. (2012). The effect of textual and interpersonal metadiscourse markers on Iranian EFL high school learners' reading comprehension. *Journal of Basic and Applied Scientific Research*, 2(6), 6210-6214.
- Bruce, R. F., & Wiebe, J. M. (1999). Recognizing subjectivity: A case study of manual tagging. *Natural Language Engineering*, 5, 187-205.
- Conrad, S., & Biber, D. (2000). Adverbial Marking of Stance in Speech and Writing. In Hunston, S., Thompson, G. (eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse* (pp. 56-73). Oxford: Oxford University Press.
- Hewings, M. (2004). "An 'important contribution' or 'tiresome reading'?" A study of evaluation in peer reviews of journal article submissions". *Journal of Applied Linguistics*, 1, 247-274.
- Hunston, S., & Sinclair, J. (2000). A local grammar of evaluation. In S. Hunston & G. Thompson (eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse* (pp. 74-101). Oxford: Oxford University Press.
- Hunston, S., & Thompson, G. (2000). (eds.) *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.
- Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, 30, 437-455.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7, 173-92.
- Kartal, G., & Sarigul, E. (2017). Frequency effects in second language acquisition: An annotated survey. *Journal of Education and Training Studies*, 6(5), 1-8. doi:10.11114/jets.v5i6.xx
- Kerbrat-Orecchioni, C. (1980). *L' énonciation. De la subjectivité dans le langage*. Paris: Armand Colin.
- Kertz, L. (2006). Evaluative adjectives: An adjunct control analysis. In D. Baumer, D. Montero, & M. Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp. 229-235). Somerville, MA: Cascadilla Proceedings Project.
- Leech, G. (1989). *An A-Z of English. Grammar & Usage*. London: Nelson.
- Marza, L. E. (2011). A comprehensive corpus-based study of the use of evaluative adjectives in promotional hotel websites. *Odisea*, 12, 97-123.
- Martin, J.R. & White, P. R. R. (2005). *The Language of Evaluation, Appraisal in English*. London and New York: Palgrave Macmillan.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (eds.). *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6-19). Cambridge: Cambridge University.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Samson, C. (2006). . . . is different from . . . : A corpus-based study of evaluative adjectives in economics discourse. *IEEE Transactions on Professional Communication*, 49(3), 236-245.
- Stotesbury, H. (2003). Evaluation in research article abstracts in the narrative and hard sciences. *Journal of English for Academic Purposes*, 2(4), 327-341.

- Swales, J. M., & Burke, A. (2003). 'It's really fascinating work': Differences in evaluative adjectives across academic registers. In P. Leistyna, & C. F. Meyer (eds.), *Corpus Analysis: Language Structure and Language Use* (pp. 1-18). New York: Rodopi.
- White, R. V. (1988). *The ELT Curriculum: Design, Innovation and Management*. Oxford: Basil Blackwell.
- Wiebe, J. (1994). Tracking point of view in narrative. *Computational Linguistics* 20(2), 233-287.
- Wiebe, J. (2000). Learning subjective adjectives from corpora. *Proceedings of 17th National Conference on Artificial Intelligence (AAAI-2000)*. Austin, Texas.
- Wiebe, J., Bruce, R., Bell, M., Martin, M., & Wilson, T. (2001). A corpus study of evaluative and speculative language. *Proceedings of the Second SIGdial Workshop on Discourse and Dialogue*. Aalborg, Denmark, 16, 1-10. doi>[10.3115/1118078.1118104](https://doi.org/10.3115/1118078.1118104)

Appendix 1. 100 most frequent adjectives in Academic texts of COCA

No	Adjectives	Frequency	No	Adjectives	Frequency	No	Adjectives	Frequency
1	Other	180834	35	Specific	28084	68	Historical	19069
2	Only	125201	36	Common	27839	69	Limited	18468
3	Social	100402	37	Available	27167	70	Primary	18301
4	Used	77200	38	Personal	25620	71	Female	18177
5	Political	69783	39	Various	25114	72	Strong	18052
6	American	67498	40	Religious	25086	73	Appropriate	17372
7	Public	61539	41	Potential	24827	74	Due	17268
8	Important	60541	42	Associated	24772	75	Negative	17260
9	National	59151	43	Positive	24340	76	Average	17258
10	United	59147	44	Academic	24098	77	Standard	17132
11	Different	58797	45	Low	23943	78	Modern	17064
12	High	55667	46	African	23508	79	Male	17051
13	Human	49509	47	Total	23399	80	Content	16584
14	Economic	48643	48	Increased	23333	81	European	16547
15	Significant	44642	49	Special	22984	82	Basic	16389
16	General	41300	50	Traditional	22800	83	Complex	15600
17	Individual	39291	51	Educational	22592	84	Moral	15459
18	Early	38933	52	Foreign	22451	85	Direct	15444
19	Given	37902	53	Natural	22348	86	Legal	15407
20	International	37380	54	Certain	21716	87	Patient	15397
21	Local	35114	55	Sexual	21660	88	Developing	15310
22	Physical	34878	56	Central	21605	89	English	15207
23	Cultural	34713	57	Effective	21514	90	Additional	15197
24	Present	34096	58	Critical	20885	91	Indian	15171
25	Environmental	33999	59	Necessary	20496	92	Ethnic	15005
26	Military	33588	60	Involved	20486	93	Main	14906
27	Likely	32768	61	Difficult	20175	94	Independent	14865
28	Possible	32758	62	Clear	19843	95	Overall	14863
29	Future	32037	63	Global	19532	96	Multiple	14848
30	Major	30800	64	Professional	19469	97	Previous	14789
31	Following	29663	65	Western	19173	98	Successful	14526
32	Particular	28644	66	Private	19109	99	Popular	14406
33	Current	28123	67	Middle	19079	100	Poor	14406
34	Similar	28108						

Appendix 2. 100 most frequent adjectives in social sciences (history, education, geography/social science, law/political science, humanities, philosophy/religion)

No	Adjective	Frequency	No	Adjective	Frequency	No	Adjective	Frequency
1	Other	133268	35	African	21061	68	European	13803
2	Only	91482	36	Foreign	20417	69	Strong	13683
3	Social	87227	37	Academic	19979	70	Ethnic	13634
4	Political	62366	38	Common	19759	71	Global	13595
5	American	56680	39	Similar	19704	72	Negative	13493
6	United	49348	40	Current	19371	73	Legal	13457
7	Used	47883	41	Sexual	18979	74	Limited	13398
8	Important	45653	42	Educational	18861	75	Moral	13395
9	Public	45598	43	Traditional	18764	76	Natural	13204
10	National	44116	44	Positive	18681	77	Appropriate	12892
11	Different	42113	45	Special	18372	78	Content	12783
12	Economic	41772	46	Various	18347	79	English	12686
13	High	37967	47	Central	16938	80	Basic	12547
14	Human	37762	48	Certain	16873	81	Chinese	12385
15	Significant	32677	49	Historical	16640	82	Key	12296
16	General	31213	50	Potential	15991	83	Popular	12188
17	Cultural	31209	51	Western	15891	84	Primary	12161
18	Individual	30706	52	Low	15734	85	Native	11956
19	International	29594	53	Critical	15697	86	Environmental	11872
20	Early	29322	54	Private	15398	87	Independent	11705
21	Physical	28504	55	Involved	15335	88	Civil	11691
22	Given	28180	56	Middle	15306	89	Due	11599
23	Present	26087	57	Necessary	15259	90	Direct	11423
24	Local	24834	58	Effective	15241	91	Written	11264
25	Likely	24099	59	Available	15223	92	Visual	11209
26	Religious	23825	60	Professional	15023	93	Poor	11073
27	Future	23362	61	Female	14925	94	Average	11004
28	Possible	22997	62	Increased	14711	95	Mental	10975
29	Particular	22469	63	Difficult	14587	96	Standard	10970
30	Military	22179	64	Total	14489	97	Main	10910
31	Major	21987	65	Modern	13891	98	Successful	10872
32	Following	21619	66	Indian	13840	99	Previous	10675
33	Specific	21266	67	Male	13813	100	Complex	10589
34	Personal	21101						

Appendix 3. 100 most frequent adjectives in science /technology and medicine

No	Adjective	Frequency	No	Adjective	Frequency	No	Adjective	Frequency
1	Other	39831	35	Common	6887	68	Initial	3960
2	Used	27347	36	Present	6869	69	Central	3938
3	Only	26331	37	American	6678	70	Complete	3924
4	Environmental	21992	38	Specific	6182	71	Relative	3913
5	High	15692	39	Economic	6088	72	Biological	3900
6	Different	14464	40	Average	5931	73	Previous	3802
7	Public	13412	41	Various	5920	74	Special	3800
8	National	13343	42	Clinical	5823	75	Strong	3694
9	Important	13152	43	Primary	5767	76	Direct	3682
10	Significant	11451	44	Effective	5742	77	Certain	3675
11	Patient	11436	45	Physical	5675	78	Technical	3626
12	Available	11017	46	Standard	5673	79	Active	3623
13	Military	10869	47	Global	5585	80	Main	3544
14	Social	10462	48	Due	5389	81	Negative	3511
15	Local	9500	49	Scientific	5291	82	Mechanical	3445
16	Human	9287	50	Additional	5281	83	Professional	3387
17	General	8845	51	Positive	5268	84	Basic	3361
18	Possible	8591	52	Particular	5067	85	Personal	3348
19	Total	8586	53	Mass	5018	86	Nuclear	3331
20	United	8445	54	Key	4994	87	Estimated	3313
21	Potential	8415	55	Political	4885	88	Traditional	3301
22	Increased	8273	56	Difficult	4792	89	Successful	3177
23	Natural	8265	57	Multiple	4768	90	Long-term	3166
24	Given	8227	58	Developing	4739	91	Useful	3107
25	Early	8056	59	Limited	4651	92	Existing	3081
26	Current	8022	60	Involved	4639	93	Original	3030
27	Major	7916	61	Necessary	4554	94	Alternative	2994
28	Likely	7818	62	Overall	4526	95	Genetic	2976
29	Low	7786	63	Complex	4510	96	Content	2934
30	Similar	7724	64	Normal	4485	97	Agricultural	2924
31	Future	7435	65	Critical	4296	98	Regional	2922
32	Individual	7402	66	Increasing	4188	99	Experimental	2910
33	Following	7205	67	Appropriate	4118	100	Industrial	2907
34	International	7174						