

A Corpus-based Analysis of English Near-synonymous Adverbs: Absolutely, Utterly

XIONG Yang-hua¹

Wuhan University of Technology, Wuhan, China

LIU Deng-feng²

Central South University, Changsha, China

The analysis of near-synonyms has always been a difficult point in English learning. There have been some relevant studies on the analysis of near-synonyms at home and abroad, but the examples are still insufficient. Based on the Corpus of Contemporary American English (COCA), this paper takes “absolutely” and “utterly” as an example, and makes a comparative analysis from the five aspects of register, collocation, colligation, semantic preference and semantic prosody, aiming to make full use of the corpus in the analysis of English near-synonyms, providing a useful supplement for second language teaching and English learning. The results show that “absolutely” is more common in spoken language while utterly is more formal; regarding the collocation, both words are more associated with adjectives and verbs; in terms of colligation patterns, “absolutely” have six kinds of grammatical patterns while there are only four grammatical patterns of “utterly”; as for semantic preference, the semantic preference of “utterly” has a wider range than that of “absolutely”; for semantic prosody, “absolutely” have more neutral semantic prosody whereas “utterly” have stronger negative semantic prosody. This case study not only provides enlightenment for English learning and teaching, but also further enriches the corpus linguistics-related research.

Keywords: corpus, near-synonymous adverbs, register, collocation, colligation, semantic preference, semantic prosody

1. Introduction

Vocabulary is an indispensable material in language communication and is the basis of language learning. As an important part of English learning, vocabulary teaching has always been of great concern to the majority of teaching researchers, especially the teaching and research of synonyms and near-synonyms. Synonyms and near-synonyms account for a large proportion of the English vocabulary, accounting for more than 60% of the total vocabulary (He Xiaodong, 2003).

Synonyms refer to words or phrases that have the same or similar meaning in the same language. Studies have shown that synonyms differ in a large number of elements, such as collocation, style, etc. (Palmer, 1981). As a result, some scholars have classified synonyms as absolute synonyms and relative synonyms. It has been

XIONG Yang-hua¹, M.A.,N/A, school of foreign languages, Wuhan University of Technology.
LIU Deng-feng², M.A.,N/A, school of foreign languages, Central South University.

found that absolute synonyms are rare (Tylor, 1995), while relative synonyms, or near synonyms, are words or phrases that have the same or similar meaning but differ in usage (Lyons, 1981).

The correct use of near-synonyms is one of the core elements of English vocabulary acquisition, and a thorough grasp of their usage can improve learners' ability to use them in practice. However, traditional methods of identifying near-synonyms in English often rely on dictionary interpretation, matching each other's meanings and explaining them with example sentences, but these methods often fail to convey the meaning of the words, and learners still use near-synonyms incorrectly in practice, thus causing misunderstandings in communication.

With the emergence of corpus linguistics, the use of corpora to distinguish between near-synonyms has become an important tool for many English language teachers and learners. The corpus provides learners with a rich corpus of language that can be better used for analysis and judgement in practice.

As a pioneer in corpus linguistics, Sinclair's (2004) lexico-grammatical theory and the framework for describing the multilayer meaning relations of lexical items have played an important role in corpus-related research. The model of extended units of meanings covers four levels of relationships: collocation, colligation, semantic preference and semantic prosody. Specifically, lexical collocation refers to the frequent collocation of different lexical items, which can be directly reflected in the text, e.g. sustainable development; class linkage refers to the grammatical characteristics of a word or a class of words, which is expressed as the combination of different lexical items in the grammatical structure and framework (Wei Naixing, 2002), e.g. *adv + v* is a class linkage. Semantic preference refers to the tendency of lexical items to be frequently co-selected semantically with surrounding words, mainly involving the semantic features or categories of habitually collocated words. Semantic prosody refers to both semantic and rhyme relationships between nodal words and their collocated lexical items.

Currently, there are more studies at home and abroad that use corpora to discriminate English near-synonyms. Most foreign studies have used corpora to compare and analyse two or more near-synonyms from the perspectives of register, lexical collocation and semantic rhyme. For example, Kennedy (2000) uses a corpus to compare and analyse the differences between the near-synonyms "between" and "through" (Biber & Conrad, 2000)

In China, most of the initial studies were qualitative, mainly involving the theories and methods of corpus linguistics on near-synonym identification, such as Li Na (2009), who studied four comparative approaches to near-synonym identification in terms of domain, collocation, semantic rhyme and class association with the help of a corpus. For example, Wang Fangfang (2018) analysed the near-sense words "acquire" and "oblige" based on the BNC corpus; Ding Chengshun (2020) took *reduce* and *decrease* as examples, and investigated the differences between them in terms of register, grammatical collocation and semantic rhyme with the help of the corpus.

However, there are relatively few empirical studies on the use of corpora for English near-sense adverbs identification. Therefore, this paper takes the pair of adverbs *absolutely* and *utterly* as a carrier and uses COCA as a research tool to analyse these pairs of adverbs in five aspects, including domain, collocation, class link, semantic preference and semantic rhyme, in order to reveal the differences between *absolutely* and *utterly* in the above five aspects and provide an empirical analysis for English teaching and learning. On the other hand, it

aims to enrich the corpus linguistics research on word sense disambiguation and provide reference for subsequent studies.

2. Research Design

2.1 Research Questions

Based on the extended meaning unit theoretical framework proposed by Sinclair, this paper aims to address the following questions.

- (1) What are the characteristics of each of Absolutely and utterly in terms of five aspects: domain, lexical collocation, class linking, semantic preference and semantic rhyme?
- (2) What are the similarities and differences in usage between Absolutely and utterly in the COCA corpus?

2.2 Research Tools

This study uses the Corpus of Contemporary American English (COCA), the world's largest free online corpus of English, which collects a total of corpus from 1990 to 2019, covering spoken language, fiction, magazines, newspapers, academic and other languages. The corpus is widely used in English teaching and research, and provides a rich corpus for linguistics, translation and other research fields.

2.3 Research Steps

Based on the COCA corpus, this study consists of five steps.

- (1) Through the CHART search function in the COCA corpus, input ABSOLUTELY and UTTERLY in turn to obtain the distribution characteristics of the two on the language domain.
- (2) Using the collocates search function in the COCA corpus, enter absolutely and utterly respectively, and set the span to -5/+5 to obtain the similarities and differences between the two collocations, and count the collocations with MI values above 3.
- (3) Using the KWIC (keywords in context) search function in the COCA corpus, 150 indexed lines were randomly examined to analyse the class link information of the two words and to conduct statistics.
- (4) To analyse the similarities and differences in the semantic tendencies of the collocations between the two, combining the high-frequency collocations and the 150 indexed rows.
- (5) Examining 150 index lines, analysing and counting the distribution of semantic rhymes between the two.

3. Research Results and Discussion

3.1 Comparative Analysis of Register Distribution

Register refers to a linguistic variant that is produced in the actual context, for communicative reasons, or because of the different professions they are engaged in, or because of the different situations, places and topics in which they speak, and is reflected in the different styles of speech, word choice and phrasing (Yang, 2002). In the COCA corpus search interface, the distribution of the adverbs absolutely and utterly in different registers was retrieved through the chart function, and the distribution is shown in Table 1.

From Table 1, it can be concluded that: firstly, the total frequency and frequency of use of absolutely is much higher than that of utterly, which can be seen from the frequency and frequency; secondly, among the eight different registers, the adverb absolutely is used most frequently in the colloquial register, indicating that

absolutely is more colloquial, while in the more formal registers, such as newspapers, novels, In contrast, utterly is used less frequently in the spoken and subtitled domains, and more frequently in novels, magazines, blogs and web pages, indicating that it is more formal than absolute.

Table 1

Word Frequency Distribution of Absolutely and Utterly in Different Domains of COCA

DOMAIN	Absolutely		Utterly	
	FREQ	PER MIL	FREQ	PER MIL
Spoken Language	38627	306.23	634	5.03
Fiction	5622	47.51	2196	18.56
Magazines	5532	43.87	1532	12.15
Newspaper	5285	43.41	740	6.08
Academic Journals	2026	16.91	897	7.49
Blog	15038	116.92	2432	18.91
Web	11767	94.7	2438	19.62
Film & Television Subtitles	14797	115.53	506	3.95

The results of the COCA corpus chart search also show the development of the words over time. The results of the search are detailed in Table 2, which shows that the development of absolutely and utterly is relatively stable on the whole, while the frequency of using absolutely fluctuates relatively more, but the general trend of utterly is decreasing. This table shows that the frequency of using absolutely and utterly fluctuates relatively more, but the overall trend of using utterly is decreasing. By counting the distribution and temporal development of near-synonyms in different language domains, we can better help learners to form “domain awareness”, so that they can use near-synonyms more accurately.

Table 2

Development Trajectory of Absolutely and Utterly over Time in COCA

TIME	Absolutely		Utterly	
	FREQ	PER MIL	FREQ	PER MIL
1990-94	10545	75.83	1139	8.19
1995-99	12321	83.38	1056	7.15
2000-04	11710	79.89	1173	8.00
2005-09	12555	86.62	1074	7.41
2010-14	12832	88.34	1052	7.24
2015-19	11926	82.40	1011	6.98

3.2 Comparative Analysis of Collocations

Partington (1998) points out from a corpus study that each word or word class has its own unique collocation behaviour. According to corpus-related theory, the higher the Mutual Information Score (MI), the higher the strength of collocation.

The Mutual Information Score is concerned with the strength of collocation between words, and refers to the comparison of the likelihood of two words occurring together with the likelihood of occurring separately. A higher MI value means that the node word has a greater influence on its lexical environment and is more attractive to its co-occurrence. Table 3 shows the collocations of the words absolutely and utterly, taking collocations with an MI value greater than 3.

From Table 3, we can see that among the top 20 collocations with an MI value greater than 3, there are 17 adjectives collocated with absolutely, accounting for 85%; there are also two verbs collocated with them, accounting for 10%, and one adverbial collocation, accounting for 5%. There are 13 adjectives with utterly (65%), 5 verbs with utterly (25%) and 2 adverbs with utterly (10%). In terms of collocations, they have the majority of adjectives in common, but a greater proportion of verbs are collocated with utterly.

Table 3

Collocations, Frequencies and MI Values for Absolutely and Utterly (span of -5/+5)

Absolutely				Utterly			
Serial Number	COLLOCATION	FREQ	MI	Serial Number	COLLOCATION	FREQ	MI
1	right	4128	3.43	1	completely	423	5.09
2	necessary	1290	3.90	2	fail	300	4.52
3	correct	933	4.18	3	destroy	279	4.91
4	essential	707	4.24	4	ridiculous	171	5.49
5	amazing	576	3.02	5	useless	124	6.37
6	ridiculous	406	3.92	6	impossible	90	3.78
7	convinced	354	4.61	7	alone	89	3.04
8	gorgeous	341	4.60	8	totally	86	3.28
9	brilliant	309	3.57	9	false	82	4.24
10	incredible	308	3.30	10	transform	72	4.45
11	fantastic	308	3.69	11	confuse	71	4.70
12	positively	297	4.96	12	incapable	68	6.69
13	stunning	285	4.70	13	meaningless	66	6.47
14	adore	214	5.08	14	devoid	65	7.57
15	delicious	211	3.76	15	lack	64	3.94
16	terrify	211	4.44	16	unable	60	4.04
17	fascinating	176	3.38	17	silent	59	4.02
18	fabulous	175	4.01	18	convinced	58	4.98
19	vital	174	3.07	19	unique	57	3.35
20	disgusting	159	3.81	20	irrelevant	56	5.51

3.3 Comparative Analysis of Class Links

Class linking refers to the “syntactic structural interrelationship between grammatical categories” (Firth, 1957). By using the corpus to distinguish the similarities and differences in the class links of near-synonyms, we can better analyse the relationships between near-synonyms from a grammatical perspective, grasp the grammatical collocation of near-synonyms, and promote the correct use of near-synonyms in practical situations. Table 4 is a statistical table produced by the author using 150 randomly selected index entries from KWIC.

From Table 4, we can see that: firstly, the types of class links are different between the two, specifically, there are more types of absolutely class links, with two more types of adv. + pron. and v. + adv. Secondly, the adv. + v. and adv. + adj. categories account for the majority of the class links, with utterly being more obvious, suggesting that adverbs often modify adjectives and verbs, a point that is also fully reflected in the previous section on word collocation. Thirdly, the v.+ adv. category accounts for a certain proportion of all links in the absolutely category, at 19.7%, indicating that absolutely is also often placed after verbs when modifying them.

In addition, it should be noted that out of the 150 randomly selected index lines, 12 index lines on absolutely are used alone, i.e. absolutely appears alone as an answer, which also confirms from the side that absolutely is more colloquial and lifelike. It can be seen that knowing the types of class linking of near-synonyms can better help learners to deepen their understanding of near-synonyms and facilitate the correct use of their grammatical collocation structures in practice.

Table 4

Comparison of Class Linkage Results Between Absolutely and Utterly in the COCA Corpus

CLASS CONNECTION	Absolutely		CLASS CONNECTION	Utterly	
	FREQ	FREQ (%)		FREQ	FREQ (%)
adv. + v.	36	27.3	adv. + v.	49	32.7
adv. + adj.	51	38.7	adv. + adj.	92	61.3
adv. + prep.	4	3.0	adv. + prep.	6	4.0
adv. + adv.	4	3.0	v.+ adv.	3	2.0
adv. + pron.	11	8.3	adv. + pron.	0	0
v.+ adv.	26	19.7	v.+ adv.	0	0
总计	132	100	总计	150	100

3.4 Comparative Semantic Preference Analysis

Semantic preference refers to the simultaneous occurrence of different lexical items with the same or similar semantic features (Paciorek & Williams, 2015). From the 20 common collocations with MI values greater than 3, it can be concluded that the collocations of absolutely and utterly are basically different. In terms of collocations, absolutely is mostly associated with amazing, gorgeous, brilliant, incredible, fantastic, positively, The words “stunning”, “amazing”, “adore” and “delicious” are all used together to express positive subjective feelings. In addition, utterly is often found alongside words such as useless, impossible and false, which indicate value judgments. This shows that although the two words are similar in meaning, there are some differences in the collocation of words, and the use of the corpus can effectively and quickly obtain statistical results to help identify the differences in semantic tendencies and avoid misuse.

3.5 Comparative Analysis of Semantic Rhymes

Firth (1957) first introduced the concept of “prosody”, which was developed by Sinclair (1991) and Louw (1993) to form the concept of “semantic rhyme”. Stubbs distinguishes between different types of semantic prosody, namely positive semantic prosody, negative semantic prosody and neutral semantic prosody. Based on the concept of semantic prosody and related descriptions, combined with contextual information and collocation information, the 150 index lines were analysed one by one to identify the different types, and the results are shown in Table 5. The difference in semantic rhyming tendencies between absolutely and utterly can help learners to better distinguish the relationship between the two, thus further increasing the correct usage rate.

Table 5

Semantic Rhyme Comparison of Absolutely and Utterly

	Positive Semantic Rhyme	Negative Semantic Rhyme	Neutral Semantic Rhyme
Absolutely	28.9%	11.1%	60%
Utterly	14%	52%	34%

Conclusion

Based on the COCA corpus, this paper analyses the near-sense adverbs absolutely and utterly in five aspects: register, word collocation, class link, semantic preference and semantic rhyme. The results show that: absolute is used in a more everyday way, while utterly is used in a more formal way; in terms of collocations, both of them are mostly used with adjectives and verbs; in terms of class links, absolute has more types of class links than utterly; in terms of semantic preferences, utterly has a wider range of semantic preferences; in terms of semantic rhymes, absolute has a neutral semantic rhyme, while utterly has a neutral semantic rhyme. In terms of semantic preference, utterly has a relatively wide range of semantic preferences; in terms of semantic rhyme, absolutely has a neutral rhyme, while utterly has a negative rhyme. This case study not only helps to identify the difference between absolutely and utterly, but also develops learners' ability to use the corpus independently and enriches the research on corpus linguistics.

References

- Biber, D., & Conrad, S. (2000). *Corpus linguistics*. Cambridge: Cambridge University Press.
- Firth, J. (1957). *Papers in linguistics*. Oxford: OUP.
- Louw, B. (1993). "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies". In M. Baker, G. Francis, & E. Togni-ni-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 157-176). Amsterdam: Benjamins. <https://doi.org/10.1075/z.64.11lou>
- Lyons, J. (1981). *Language and linguistics*. Cambridge University Press, Cambridge. M.
- Paciorek, A., & Williams, J. N. (2015). Implicit learning of semantic preferences of verbs. *Studies in Second Language Acquisition*, (37), 359-382.
- Palmer, F. R. (1981). *Semantics*. Cambridge: Cambridge University Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge Kennedy.
- Taylor, J. R. (1995). *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Oxford University Press.
- 丁成顺. (2020). 基于COCA语料库的近义词研究——以“reduce”和“decrease”为例. *汉字文化*, (7).
- 贺晓东. (2003). *英语同义词词典*. 北京:商务印书馆.
- 李娜, 陈德生. (2009). 基于语料库的英语同义词辨析教学方法探究. *教育探索*.
- 王芳芳. (2018). 基于 BNC 语料库的近义词用法对比研究——以“acquire”和“obtain”为例. *海外英语*, (21), 100.
- 王乃兴. (2002). 语义韵研究的一般方法. *外语教学与研究*, (4), 80-81.
- 杨慧中. (2002). *语料库语言学导论*. 上海: 上海外语出版社.