РОЗДІЛ 6
СТРУКТУРНА, ПРИКЛАДНА ТА МАТЕМАТИЧНА ЛІНГВІСТИКА

CORPUS-BASED STUDY OF COLLOCATIONS WITH SYNONYMS 'FREEDOM' AND 'LIBERTY' IN AMERICAN ENGLISH

КОРПУСНО-БАЗОВАНЕ ДОСЛІДЖЕННЯ КОЛОКАЦІЙ З СИНОНІМАМИ 'FREEDOM' ТА 'LIBERTY' В АМЕРИКАНСЬКІЙ АНГЛІЙСЬКІЙ МОВІ

Bihdai M.O.,
orcid.org/ 0000-0001-6507-6100
Candidate of Philological Sciences,
Lecturer at the Applied Linguistics Department
Lviv Polytechnic National University

The purpose of the article is to conduct the corpus-based research of collocations with synonyms 'freedom' and 'liberty' in American English based on the Corpus of Contemporary American English (COCA). The topicality of this study is determined by the need to conduct linguistic research based on American English on a large array of data.

The COCA comprises approximately 1 billion words spanning from 1990 to 2019, with an even distribution of genres including spoken language, fiction, popular magazines, newspapers, academic journals, television/film subtitles, blogs, and web pages. Each genre contains roughly 120–130 million words, making the COCA a unique resource for linguistic research.

Analysis of the collocations of 'freedom' and 'liberty' with nouns, verbs in infinitive, and adjectives identified both shared and distinct characteristics between them. The presented data showed that there is 81% higher frequency in the usage of the word 'freedom' compared to 'liberty' in American English.

Upon analyzing the collocations 'freedom + noun' and 'liberty + noun', it was observed that the only noun that features in their top 10 collocations is 'movement'. Additionally, the 'liberty movement' appears more frequently than the 'freedom movement' in American English. Furthermore, comparing the collocations 'adjective + freedom' and 'adjective + liberty' showed that the adjectives 'religious', 'individual', 'personal', 'economic', 'political', and 'human' are common in both synonymous pairs.

Synonyms 'freedom' and 'liberty' have the biggest number of similarities when forming collocations with adjectives and the least similarities were observed with collocations with nouns in American English based on data from COCA during the period 1990–2019. We see the prospect of further research in the study of collocations with these synonyms across different subgenres based on the data of this corpus.

Key words: corpus linguistics, corpus, corpus-based approach, collocations, synonyms, American English.

Мета статті – провести корпусно-базоване дослідження колокацій із синонімами 'freedom' і 'liberty' в американській англійській мові на основі Корпусу сучасної американської англійської мови (COCA). Актуальність даної розвідки зумовлена необхідністю проведення лінгвістичних досліджень на основі американської англійської мови на великому масиві даних.

COCA містить приблизно 1 мільярд слів, що охоплюють період з 1990 по 2019 роки, а також рівномірний розподіл жанрів, включаючи розмовну мову, художню літературу, популярні журнали, газети, наукові журнали, субтитри до телебачення та фільмів, блоги та веб-сторінки. Кожен жанр складається з приблизно 120-130 мільйонів слів, що робить COCA унікальним ресурсом для лінгвістичних досліджень.

Аналіз колокацій 'freedom' і 'liberty' з іменниками, дієсловами в інфінітиві та прикметниками виявив їхні спільні та відмінні риси. Представлені дані показали, що в американській англійській мові слово 'freedom' вживається на 81% частіше, ніж 'liberty'.

Проаналізувавши словосполучення 'freedom + іменник' і 'liberty + іменник', виявлено, що 'movement' – це єдиний іменник, який є в топ-10 колокаціях для обох слів. Крім того, в американській англійській мові словосполучення 'liberty movement' з'являється частіше, ніж 'freedom movement'. Порівняння колокацій 'прикметник + freedom' і 'прикменик + liberty' показало, що прикметники 'religious', 'individual', 'personal', 'economic', 'political' та 'human' є спільними для обох синонімічних пар.

Синоніми 'freedom' і 'liberty' найбільше схожі при формуванні колокацій з прикметниками, а найменше – у колокаціях з іменниками в американській англійській мові на основі даних COCA за період 1990-2019 років. Перспективу подальших розвідок вбачаємо у вивченні колокацій з цими синонімами в різних піджанрах на основі даних цього корпусу.

Ключові слова: корпусна лінгвістика, корпус, корпусно-базований підхід, колокації, синоніми, американська англійська мова.

Formulation of the problem. Corpus linguistics is one of the priority areas of modern linguistics. It involves the study of large, structured collections of naturally occurring language data, called corpora. The history of corpus linguistics dates back to the early 1960s when computational linguists began exploring ways to use computers to analyze and study a language.

The development of corpus linguistics is closely related to the development of computer technologies. With the advent of more powerful computers and advancements in natural language processing techniques, corpus linguistics has become increasingly popular and sophisticated. Researchers use corpora to investigate a wide range of linguistic phenomena, including syntax, semantics, pragmatics, discourse analysis, and language variation and change.

Today, corpus linguistics is used in various fields, such as language teaching and learning, lexicography, forensic linguistics, and language technology. The development of digital corpora and analysis tools has also made corpus linguistics a more accessible and user-friendly research approach, allowing for more detailed and nuanced analyses of language data.

Corpus studies make it possible to objectively study various linguistic phenomena since they are carried out on the materials of natural language texts.

The **topicality** of the study is determined by the need to conduct linguistic research based on American English on a large array of data.

The **purpose** of the article is to conduct the corpus-based research of collocations with synonyms 'freedom' and 'liberty' in American English based on the Corpus of Contemporary American English (COCA).

Analysis of recent research and publications. Corpus-based research has become a crucial aspect of contemporary linguistics, with the application of diverse methods and approaches allowing for increasingly complex explorations. For over three decades, English language corpora such as the American National Corpus, Corpus of Contemporary American English, and the Brown Corpus have served as foundational resources for numerous research worldwide.

Currently, a large number of studies are carried out in the field of corpus linguistics for English [1, 3, 5, 6, 10, 11, 14], Ukrainian [8, 9], and other languages. However, there are not enough studies concerned with the analysis of collocations with synonyms 'freedom' and 'liberty' based on American English and COCA.

A corpus is a large and structured collection of texts or spoken language data that is used to study the language. A corpus may consist of written texts, such as novels, articles, or legal documents, or spoken language data, such as recordings of conversations or speeches. The texts or recordings in a corpus are often selected to represent a particular language, genre, or time period, and they are typically annotated with information about their structure, vocabulary, and other linguistic features. Linguists use corpora to study language in a data-driven way, analyzing patterns of language use and exploring how language varies across different contexts and speakers.

The term "corpus linguistics" was first mentioned in 1983 in the collection of scientific works "Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research". The researchers D. Biber [2], W. Francis [4], G. Leech [7], and J. Sinclair [12] made a big impact on corpus linguistics and corpus studies all over the world.

By examining the collocations or nearby words of a word or phrase, it is possible to gain valuable insight into its meaning and usage. This is due to the idea that you can learn a lot about a word by the words it frequently appears with. With COCA, a genre-balanced corpus of one billion words, users can observe the frequency of collocations by part of speech. This provides information on whether the collocations tend to appear before or after the word in question and how closely they are associated.

"Researchers differ with respect to what types of co-occurrence they focus on when identifying collocations. Some treat co-occurrence as a purely sequential phenomenon defining collocates as words that co-occur more frequently than expected within a given span. Some researchers require a span of 1 (i.e., the words must occur directly next to each other), but many allow larger spans (five words being a relatively typical span size) [13, p. 220]". In our study of collocations with synonyms 'freedom' and 'liberty' and nouns, verbs in the infinitive form, and adjectives only the words that occurred directly next to each other were investigated.

The presented research is corpus-based and implies a comparison method, as well as frequency analysis. Our study is based on data from the COCA corpus [https://www.english-corpora.org/coca/] as one of the best genre-balanced corpora.

COCA was selected as the source for materials, as it is a valuable resource for linguists and other researchers interested in the study of American English. Its size, variety of sources, and searchability make it a powerful tool for investigating a wide range of linguistic phenomena, and its time period and part-of-speech tagging enable researchers to explore language use in a contemporary context.

Results and discussions. "Natural languages typically contain pairs (or larger sets) of words with very similar meanings, such as big and large, begin and start or high and tall. In isolation, it is often difficult to tell what the difference in meaning is, especially since they are often interchangeable at least in some contexts. Obviously, the distribution of such pairs or sets with respect to other words in a corpus can provide insights into their similarities and differences" [13, p. 235–236].

Taking this into account the presented research shows similarities and differences between the synonyms 'freedom' and 'liberty' in general in American English across genres and time periods, as well as depicts how these synonyms form collocations with nouns, verbs in infinitive, and adjectives.

Fig. 1 displays the distribution of the word 'freedom' between the years 1990 to 2019. The highest frequency of 10639 was observed during the period 1990–1994. Subsequently, a gradual reduction in frequency occurred from 9593 during 199–1999, to 8521 during 2000–2004, and finally to 7528 during 2005–2009. A subsequent increase in frequency was observed between 2010–2014 with a frequency of 7798 and continued up to 2015–2019 with a frequency of 7980. However, this increase was still sig-

nificantly lower than the frequency observed during 1990–1994.

Fig. 2 illustrates the frequency of the word 'liberty' from 1990 to 2019. Similar to the findings for 'freedom', the highest frequency of 2965 was observed in texts from 1990–1994. This was followed by a gradual reduction in frequency from 2813 during 1995–1999, to 2568 in 2000–2004, 2573 in 2005–2009, and 2530 in 2010–2014. Notably, a subsequent increase in frequency was observed from 2015-2019 with a frequency of 2845, following a similar trend as observed for 'freedom'.

Fig. 3 presents a graphical representation of the word 'freedom' across various genres. Academic journals exhibited the highest frequency with 15624 occurrences, followed by web pages and blogs with similar frequencies of 14907 and 14856, respectively. The frequency of 'freedom' in popular magazines, spoken texts, and newspapers was 9285, 9076, and 8706, respectively. Television/film subtitles exhibited the least frequency with 5251, and fiction had the lowest frequency with 4117.

Fig. 4 displays the frequency of the word 'liberty' across diverse genres. Similar to 'freedom', the highest frequency of 4653 occurrences for 'liberty' is observed in academic journals, followed by newspapers with

SECTION	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	10639	9593	8521	7528	7798	7980
WORDS (M)	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	87.85	76.62	68.37	61.18	63.22	65.01
SEE ALL SUB-SECTIONS AT ONCE						

Fig. 1. Frequency of the word 'freedom' across time periods

SECTION	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	2965	2813	2568	2573	2530	2845
WORDS (M)	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	24.48	22.47	20.61	20.91	20.51	23.18
SEE ALL SUB-SECTIONS AT ONCE						

Fig. 2. Frequency of the word 'liberty' across time periods

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD
FREQ	81822	14856	14907	5251	9076	4117	9285	8706	15624
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8
PER MIL	82.40	115.51	119.97	41.00	71.95	34.79	73.64	71.51	130.43
SEE ALL SUB-SECTIONS AT ONCE									

Fig. 3. Frequency of the word 'freedom' across different genres

3220 and popular magazines with 2807, whereas spoken texts exhibit 2100, television/film subtitles with 1855, and fiction with 1659. In contrast to 'freedom', the lowest frequency for 'liberty' is observed in blogs and web pages. In general, the presented data reveal that 'freedom' is 81% more frequently used in American English than 'liberty'.

Upon closer examination of the frequency of the word 'freedom' within the subgenres of the genre of the academic journal (Fig. 5), it is evident that the three most prominent subgenres are law and political science with a frequency of 3136, philosophy and religion with a frequency of 2664, and history with a frequency of 2590.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD
FREQ	16294	0	0	1855	2100	1659	2807	3220	4653
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8
PER MIL	16.41	0.00	0.00	14.48	16.65	14.02	22.26	26.45	38.84
SEE ALL SUB-SECTIONS AT ONCE									

Fig. 4. Frequency of the word 'liberty' across different genres

SECTION	History	Education	Geog/SocSci	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine	Misc	Business
FREQ	2590	1356	1972	3136	1685	2664	618	207	1310	63
WORDS (M)	13.4	15.8	20.0	12.3	16.2	7.8	17.5	10.8	4.8	1.2
PER MIL	193.35	85.96	98.46	255.26	103.99	339.73	35.41	19.15	271.91	53.38
CLICK FOR CONTEXT										

Fig. 5. Frequency of the word 'freedom' across different subgenres of academic genre

SECTION	History	Education	Geog/SocSci	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine	Misc	Business
FREQ	631	96	505	1855	409	606	80	50	380	37
WORDS (M)	13.4	15.8	20.0	12.3	16.2	7.8	17.5	10.8	4.8	1.2
PER MIL	47.11	6.09	25.21	150.99	25.24	77.28	4.58	4.63	78.88	31.35
CLICK FOR CONTEXT										

Fig. 6. Frequency of the word 'liberty' across different subgenres of academic genre

Fig. 6 illustrates the frequency of the word 'liberty' in various subgenres of the academic journal genre. The subgenre with the highest frequency of 1855 is law and political science. However, in contrast to the word 'freedom', the frequency of 'liberty' is notably lower in the subgenres of history, philosophy, and religion.

Collocations can be used to distinguish between words with similar meanings. Additionally, the meaning and usage of a word can vary depending on the genre in which it appears. With the help of collocations, it is also possible to examine changes in the meaning of words over time. Researchers can analyze collocations from different time periods to identify shifts in meaning or usage. Typically, collocations look at words that appear within a four-word range to the left or right of the target word, whereas in our study one-word range was investigated.

In our study, an analysis of the collocations of 'freedom' and 'liberty' with nouns, verbs in infinitive, and adjectives has been conducted which produced the below-mentioned results.

Fig. 7 displays the collocation of 'freedom + noun' across different genres and time periods. The top 10 most frequent collocations are: 'freedom fighters', 'freedom house', 'freedom fighter', 'freedom press', 'freedom party', 'freedom caucus', 'freedom agenda', 'freedom restoration', 'freedom forum', and 'freedom movement'. The majority of these collocations occur most frequently in academic journals, television/film subtitles, newspapers, and spoken genres. Over time, there has been a shift in the frequency distribution of these collocations, from 'freedom fighter(s)', 'freedom party', and 'freedom

movement' in the 1990–1994 period to 'freedom house' and 'freedom agenda' in 2010–2014, and finally to 'freedom caucus' in 2015–2019.

Fig. 8 displays the frequency of the collocation of 'liberty + noun' across different genres and time periods. The top 10 most frequent collocations are: 'liberty university', 'liberty media', 'liberty lady', 'liberty radio', 'liberty city', 'liberty county', 'liberty reserve', 'liberty movement', 'liberty ships', and 'liberty bowl'. The majority of these collocations have the highest occurrence in genres such as newspapers, spoken, popular magazines, and blogs. The distribution of frequency within time periods shows that the collocations of 'liberty' shifted from 'liberty radio' in 1990–1994 to 'liberty city' and 'liberty ships' in 2005–2009 to 'liberty lady' and 'liberty reserve' in 2015–2019.

Upon comparing the collocations 'freedom + noun' and 'liberty + noun', it can be observed that the sole noun they have in common among their top 10 collocations is 'movement'. Additionally, it can be noted that 'liberty movement' has a higher frequency than 'freedom movement' in American English.

Fig. 9 illustrates the distribution of the collocation 'verb in infinitive + freedom' across various genres and time periods. The top 10 most frequent collocations observed are: 'have freedom', 'be freedom', 'defend freedom', 'protect freedom', 'bring freedom', 'find freedom', 'enjoy freedom', 'promote freedom', 'mean freedom', and 'want freedom'. The majority of these collocations are commonly found in blogs, web pages, and spoken genres. These top 10 collocations had the highest frequency during 1990–1994, and a decreasing trend can be observed until 2015–2019.

HELP.	①	★		ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19	SEC1
1	⓪	★	FIGHTERS	631	90	101	117	86	33	65	68	71	133	68	99	49	52	41	631
2	⓪	★	HOUSE	370	25	32	1	62		51	63	136	24	18	30	42	113	86	370
3	⓪	★	FIGHTER	342	49	36	66	47	23	36	47	38	72	38	50	29	33	35	342
4	⓪	★	PRESS	303	30	45	2	39	1	37	81	68	49	25	50	35	23	40	303
5	⓪	★	PARTY	301	34	16	6	92	4	42	77	30	104	40	53	8	13	33	301
6	⓪	★	CAUCUS	237				164		37	36							237	237
7	⓪	★	AGENDA	140	27	9		20		9	6	69				15	89		140
8	⓪	★	RESTORATION	137	11	6		6		22	25	67	3	48	5	1	21	42	137
9	⓪	★	FORUM	136	1	15		20		4	85	11	48	37	32		3		136
10	⓪	★	MOVEMENT	132	23	28	13	7	1	30	15	15	25	21	10	8	10	7	132
11	⓪	★	FOUNDATION	124	15	25		24	6	21	27	6	11	33	14	7	7	12	124
12	⓪	★	RING	113	21	25	25	17	5	13	5	2	11	11	21	9	13	2	113
13	⓪	★	RIDERS	108	2	9	11	24	13	29	11	9	18	16	17	10	13	23	108
14	⓪	★	SUMMER	108	1	7	4	35	4	24	20	13	41	6	11	15	17	10	108
15	⓪	★	ACT	103	5	11		8	2	27	15	35	13	14	8	13	7	32	103
16	⓪	★	INTERNET	98	26	43	2	3		9	4	11				3	14	12	98
17	⓪	★	COALITION	91	8	21	3	30	1	8	18	2	4	1	5	3	35	14	91
18	⓪	★	TOWER	91	12	12	3	20	1	10	33				10	37	7	13	91
19	⓪	★	STRUGGLE	71	4	7	2	14		16	4	24	9	16	10	10	5	10	71
20	⓪	★	TRAIL	69	13	3	12	2	3	16	20		4	15	7	3	18	0	69

Fig. 7. Frequency of the collocation 'freedom + noun' across different genres and time periods

ON CLICK: CONTEXT TRANSLATE (??) ENTIRE PAGE GOOGLE IMAGE PRON/VIDEO BOOK (HELP)

HELP	(i)	★		ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19	SEC1
1	0	★	UNIVERSITY	266	37	44	3	90	2	41	33	16	12	16	17	43	38	59	266
2	0	★	MEDIA	209	33	17		7		23	127	2	25	30	36	37	23	8	209
3	0	★	LADY	190	32	24	23	39	18	26	26	7	10	10	22	18	39	35	190
4	0	★	RADIO	127	27	10		22	8	30	21	9	57	4	10	3	10	6	127
5	0	★	CITY	107	16	14	22	12	3	16	20	4	15	11	7	21	8	15	107
6	0	★	COUNTY	104	1	3		1	1	4	88	6	4	22	21	22	30	1	104
7	0	★	RESERVE	92	3					87	2						1	88	92
8	0	★	MOVEMENT	88	58	21	1	2		1	5		3			4	2		88
9	0	★	SHIPS	87	9	3	1		4	21	2	47	4	2	2	56	4	7	87
10	0	★	BOWL	74	6	4		1	1	8	54		11	29	11	4	6	3	74
11	0	★	ISLAND	70	12	10	11	17	6	4	8	2		5	16	2	14	11	70
12	0	★	SHIP	70	6	3	14		7	14	10	16	18	4	4	27	2	6	70
13	0	★	USS	62	36	18	6				1	1		1		6		1	62
14	0	★	SQUARE	57	5	44		1	5	2			1		5		1		57
15	0	★	STREET	55	8	9	5	4	5	7	13	4	3	4	17	5	2	7	55
16	0	★	INTEREST	54	4	9		7		1	3	30	6	5	12	3	10	5	54
17	0	★	VALANCE	52	8	8	6	3	1	5	7	16	6	18	2		10	2	52
18	0	★	AVENUE	47	4	4	22		5	1	11			1	21	2	6	9	47
19	0	★	HEIGHTS	46	5	2	3	14		1	21		1	19	18		1		46
20	0	★	HILL	43	1	5		2		6	27	2	5	2	1	4	20	5	43

Fig. 8. Frequency of the collocation 'liberty + noun' across different genres and time periods

ON CLICK: CONTEXT TRANSLATE (??) ENTIRE PAGE GOOGLE IMAGE PRON/VIDEO BOOK (HELP)

HELP	(i)	★		ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19	SEC1
1	0	★	HAVE	267	59	67	18	68	6	9	22	18	29	22	22	14	31	13	267
2	0	★	BE	65	19	9	5	13	4	3	9	3	3	9	10	7	3	5	65
3	0	★	DEFEND	62	8	16	2	15	1	11	5	4	9	4	5	8	8	4	62
4	0	★	PROTECT	49	8	10	4	9	2	3	3	10	14	3	1	4	3	6	49
5	0	★	BRING	48	10	3	9	10	2	5	5	4	9	4	6	8	5	3	48
6	0	★	FIND	37	11	8	1	8	5	2	1	1	2	1	6	4	1	4	37
7	0	★	ENJOY	34	6	4	1	5	1	2	3	12	6	4	4	4	2	4	34
8	0	★	PROMOTE	33	2	10	1	4		6	4	6	2	2	5	3	6	3	33
9	0	★	MEAN	31	5	5	1	7	2	4	3	4	8	2	3	4	2	2	31
10	0	★	WANT	30	11	7	2	7	1	2			2	1	3	3	3		30
11	0	★	PRESERVE	29	5	7	2	3	3		2	7	2	4	3	4	1	3	29
12	0	★	ACHIEVE	28	6	4		5		3	2	8	4	3	4	3	1	3	28
13	0	★	GAIN	28	5	10	2	3	1	1	5	1	3	2	3		3	3	28
14	0	★	ALLOW	27	7	8	1	5		2		4	2	2	1	6		1	27
15	0	★	GUARANTEE	27	3	3		3		6	6	6	3	3	5	1	5	4	27
16	0	★	SUPPORT	26	6	9		4		1	2	4	3	2	1	1	3	1	26
17	0	★	SEE	25	2	2	7	10	1	3			6	1	5	1	5	3	25
18	0	★	RESTORE	22	7	5		5		1	1	2	3	3	2	1		1	22
19	0	★	RESTRICT	22	5	7		4		1	2	3	2	3	1	2			22
20	0	★	SEEK	21	4	3		5	2	2	3	1	4	2	2	1	3	2	21

Fig. 9. Frequency of the collocation 'verb in infinitive + freedom' across different genres and time periods

The collocation 'verb in infinitive + liberty' across different genres and time periods is displayed on Fig. 10. The top 10 most frequent collocations are: 'have liberty', 'protect liberty', 'preserve liberty', 'proclaim liberty', 'defend liberty', 'secure liberty', 'bring liberty', 'promote liberty', 'restore liberty', and 'advance liberty'. The majority of these collocations occur most frequently in blogs, web pages, and academic journals.

Upon comparing the collocations 'verb in infinitive + freedom' and 'verb in infinitive + liberty', it is observed that the verbs 'have', 'protect', 'defend', and 'promote' are shared between these synonymous pairs.

Fig. 11 presents the collocation analysis of 'adjective + freedom' across diverse genres and timeframes. The top ten frequently occurring collocations are 'religious freedom', 'academic freedom', 'individual freedom', 'personal freedom', 'economic freedom', 'political freedom', 'intellectual freedom', 'human freedom', 'greater freedom', and 'Iraqi freedom'. The majority of these collocations appear in academic journals, blogs, and web pages. The distribution of collocations has undergone a significant change over time, with a shift from 'academic freedom', 'individual freedom', and 'human freedom' in the 1990–1994 period to 'religious freedom', 'personal freedom', and 'economic freedom' in 1995–1999. During 2000–2009, 'Iraqi freedom' was the most
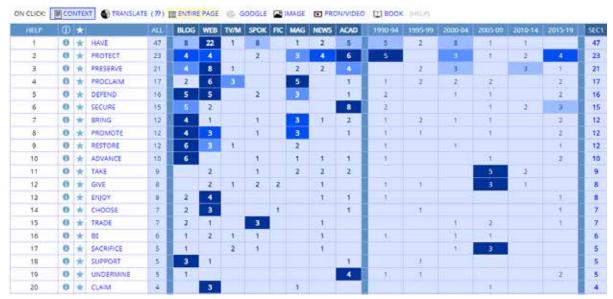
ON CLICK: CONTEXT · TRANSLATE (77) · ENTIRE PAGE · GOOGLE · IMAGE · PRON/VIDEO · BOOK (HELP)

HELP				ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19	SEC1
1	ⓘ	★	HAVE	47	8	22	1	8		1	2	5	5	2	8	1	1		47
2	ⓘ	★	PROTECT	23	4	4		2		3	4	6	5		3	1	2	4	23
3	ⓘ	★	PRESERVE	21	4	8	1			2	2	4		2	3		3	1	21
4	ⓘ	★	PROCLAIM	17	2	6	3			5		1	1	2	2	2		2	17
5	ⓘ	★	DEFEND	16	5	5		2		3		1	2	1	1			2	16
6	ⓘ	★	SECURE	15	5	2						8	2			2	3		15
7	ⓘ	★	BRING	12	4	1		1		3	1	2	1	2		1		2	12
8	ⓘ	★	PROMOTE	12	4	3		1		3		1	1	1		1		2	12
9	ⓘ	★	RESTORE	12	6	3	1			2				1				1	12
10	ⓘ	★	ADVANCE	10	6			1		1	1	1				1		2	10
11	ⓘ	★	TAKE	9		2		1		2	2	2				5	2		9
12	ⓘ	★	GIVE	8		2	1	2	2		1		1	1		3	1		8
13	ⓘ	★	ENJOY	8	2	4					1	1	1					1	8
14	ⓘ	★	CHOOSE	7	2	3			1			1		1				1	7
15	ⓘ	★	TRADE	7	2	1		3			1				1	2		1	7
16	ⓘ	★	BE	6	1	2	1	1				1	1		1	1			6
17	ⓘ	★	SACRIFICE	5	1			2	1		1					1	3		5
18	ⓘ	★	SUPPORT	5	3	1						1		1					5
19	ⓘ	★	UNDERMINE	5	1							4	1	1				2	5
20	ⓘ	★	CLAIM	4		3			1							1			4

Fig. 10. Frequency of the collocation 'verb in infinitive + liberty' across different genres and time periods

frequently occurring collocation, followed by 'intellectual freedom' in 2015–2019.

Fig. 12 illustrates the collocation of 'adjective + liberty' across various genres and time periods. The top ten frequently occurring collocations are: 'religious liberty', 'individual liberty', 'personal liberty', 'civil liberty', 'economic liberty', 'human liberty', 'American liberty', 'political liberty', 'negative liberty', and 'ordered liberty'. Same as for 'freedom', the majority of collocations with 'liberty' occur most frequently in academic journals, blogs, and web pages. The distribution of collocations has undergone a significant change over time, with a shift from 'personal liberty', 'human liberty', and 'political liberty' in the 1990–1994 period to 'individual liberty' and 'negative liberty' in 1995–1999, and finally to

'religious liberty', 'civil liberty', 'economic liberty', and 'American liberty' in 2015–2019.

A comparison of the collocations 'adjective + freedom' and 'adjective + liberty' reveals that the adjectives 'religious', 'individual', 'personal', 'economic', 'political', and 'human' are common for both of these synonymous expressions.

One of the most interesting functions of the COCA is to compare collocations, which can be especially useful when studying synonyms. Fig. 13 shows such a comparison for the phrases 'adjective + freedom' and 'adjective + liberty'. Both tables are sorted by the frequency of occurrence of each adjective with the studied nouns 'freedom' (W1) and 'liberty' (W2). The overall ratio W1/W2 for 'freedom'/'liberty' is 3. For each collocation the ratio W1/W2 is shown,
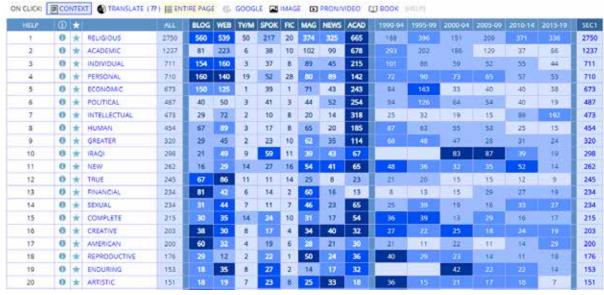
ON CLICK: CONTEXT · TRANSLATE (77) · ENTIRE PAGE · GOOGLE · IMAGE · PRON/VIDEO · BOOK (HELP)

HELP				ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19	SEC1
1	ⓘ	★	RELIGIOUS	2750	560	539	50	217	20	374	325	665	188	396	151	209	371	336	2750
2	ⓘ	★	ACADEMIC	1237	81	223	6	38	10	102	99	678	293	202	186	129	37	86	1237
3	ⓘ	★	INDIVIDUAL	711	154	160	3	37	8	89	45	215	101	86	59	52	55	44	711
4	ⓘ	★	PERSONAL	710	160	140	19	52	28	80	89	142	72	90	73	65	57	53	710
5	ⓘ	★	ECONOMIC	673	150	125	1	39	1	71	43	243	84	163	33	40	40	38	673
6	ⓘ	★	POLITICAL	487	40	50	3	41	3	44	52	254	94	126	64	54	40	19	487
7	ⓘ	★	INTELLECTUAL	473	29	72	2	10	8	20	14	318	25	32	19	15	89	193	473
8	ⓘ	★	HUMAN	454	67	89	3	17	8	65	20	185	87	63	55	53	25	15	454
9	ⓘ	★	GREATER	320	29	45	2	23	10	62	35	114	68	48	47	28	31	24	320
10	ⓘ	★	IRAQI	298	21	49	9	59	11	39	43	67			83	87	39	19	298
11	ⓘ	★	NEW	262	16	29	14	27	16	54	41	65	48	36	32	35	52	14	262
12	ⓘ	★	TRUE	245	67	86	11	11	14	25	8	23	21	20	15	15	12	9	245
13	ⓘ	★	FINANCIAL	234	81	42	6	14	2	60	16	13	8	13	15	29	27	19	234
14	ⓘ	★	SEXUAL	234	31	44	7	11	7	46	23	65	25	39	19	16	33	27	234
15	ⓘ	★	COMPLETE	215	30	35	14	24	10	31	17	54	36	39	13	29	16	17	215
16	ⓘ	★	CREATIVE	203	38	30	8	17	4	34	40	32	27	22	25	18	24	19	203
17	ⓘ	★	AMERICAN	200	60	32	4	19	6	28	21	30	21	11	22	11	14	29	200
18	ⓘ	★	REPRODUCTIVE	176	29	12	3	22	1	50	24	36	40	29	23	14	11	18	176
19	ⓘ	★	ENDURING	153	18	35	8	27	2	14	17	32			42	22	22	14	153
20	ⓘ	★	ARTISTIC	151	18	19	7	23	8	25	33	18	36	15	21	17	18	7	151

Fig. 11. Frequency of the collocation 'adjective + freedom' across different genres and time periods

Fig. 12. Frequency of the collocation 'adjective + liberty' across different genres and time periods

Fig. 13. Comparison of the collocations 'adjective + freedom' and 'adjective + liberty' (sorted by frequency)

for example, there are 47.3 times as many cases of 'intellectual freedom' as there are of 'intellectual liberty'. The score shows the ratio of W1/W2 for this collocation to overall W1/W2 for this word. In the case of 'intellectual freedom', the score is 15.8.

Conclusions. The presented data indicate that the word 'freedom' is used 81% more frequently in American English than 'liberty'. Furthermore, the highest frequency was observed in texts dated between 1990-1994 both for 'freedom' and 'liberty'.

Upon comparing the collocations of the synonymous pairs 'verb in infinitive + freedom' and 'verb in infinitive + liberty', it was observed that the verbs 'have', 'protect', 'defend', and 'promote' are common between them.

Similarly, upon analyzing the collocations 'freedom + noun' and 'liberty + noun', it can be observed that the only noun that features in their top 10 collocations is 'movement'. Additionally, the 'liberty movement' appears more frequently than the 'freedom movement' in American English.

Furthermore, comparing the collocations 'adjective + freedom' and 'adjective + liberty' highlights that the adjectives 'religious', 'individual', 'personal', 'economic', 'political', and 'human' are common in both synonymous pairs.

Thus, synonyms 'freedom' and 'liberty' have the biggest number of similarities when forming collocations with adjectives and the least similarities were observed with collocations with nouns in American English based on data from COCA during 1990–2019. We see the prospect of further research in the study of collocations with these synonyms across different subgenres based on the data of this corpus.

REFERENCES:

1. Arizavi S., Choubsaz Y. To Use or Not to Use the Shorter Forms: A Corpus-Based Analysis of the Apologetic Expressions "Sorry and I'm sorry" in American Spoken English Discourse. *Corpus Pragmatics.* 2019. Vol. 3. P. 21–47.

2. Biber D. Using corpus-based methods to investigate grammar and use: some case studies on the use of verbs in English. 2001. P. 101–115.

3. Bihdai M., Kalymon Y., Syntactic concepts analyzer based on the English complex sentences with an object clause. *Proceedings of the 6th International Conference on Computational Linguistics and Intelligent Systems COLINS 2022.* Gliwice, Poland. 2022. Vol. 1. P. 751–762.

4. Francis W. Language Corpora. Berlin – New York: Mouton de Gruyter. 1991. P. 17–35.

5. Jucker A. Apologies in the History of English: Evidence from the Corpus of Historical American English (COHA). *Corpus Pragmatics.* 2018. Vol. 2. P. 375–398.

6. 6. Kondo Y. A Corpus-Based Study of the Concept of 'Luxury' Using Web-Crawled Corpora, enTenTen 2013 and ukWaC. *Corpus Pragmatics.* 2019. Vol. 3. P. 1–20.

7. Leech G. New resources, or just better old ones? Corpus Linguistics and the Web. Amsterdam: Rodopi. 2007. P. 134–149.

8. Levchenko O., Dilai M. A Method of Automated Corpus-Based Identification of Metaphors for Compiling a Dictionary of Metaphors: A Case Study of the Emotion Conceptual Domain. *IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT).* 2021. P. 52–55.

9. Levchenko O., Tyshchenko O., Dilai M. Automated Identification of Metaphors in Annotated Corpus (Based on Substance Terms). *Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems. COLINS 2021*. Lviv, Ukraine. 2021. Vol. 1. P. 16–31.

10. Peppard J. BE Like Simile Usage Across Genres in the COCA. *Corpus Pragmatics*. 2023. P. 1–26.

11. Rafatbakhsh E., Ahmadi A. A thematic corpus-based study of idioms in the Corpus of Contemporary American English. *Asian-Pacific Journal of Second and Foreign Language Education*. 2019. Vol. 4. P. 1–21.

12. Sinclair J. Corpus, Concordance, Collocation Oxford: Oxford University Press. 1991. 170 p.

13. Stefanowitsch A. Corpus Linguistics: A Guide to the Methodology. Berlin: Language Science Press. 2020. 510 p.

14. Whitty L. A Reanalysis of the Uses of Can and Could: A Corpus-Based Approach. *Corpus Pragmatics.* 2019. Vol. 3. P. 225–247.