The Usage of Because of-Words in British National Corpus

Bei Yang

School of English and Education, Guangdong University of Foreign Studies, China

Abstract

Synonymy is an important yet intricate linguistic feature in the field of lexical semantics. Using the 100 million-word British National Corpus (BNC) as data and the software Sketch Engine (SkE) as an analyzing tool, this paper explores the collocational behavior and semantic prosodies of near synonyms *in virtue of*, *owing to*, *thanks to*, *as a result of*, *due to* and *because of*. The results show that these near synonyms differ in their collocational behavior and semantic prosodies. The pedagogical implications of the findings are also discussed.

Keywords: Collocation; Semantic prosody; Near synonym.

1. Introduction

Synonymy, or semantic equivalence, is an important yet intricate linguistic feature in the field of lexical semantics. In this paper, synonyms refer to lexical pairs that have very similar cognitive meanings or denotational meanings, but which may differ in collocations or semantic prosodies. Any natural language consists of a considerable number of synonymous words. English is particularly rich in synonyms because of historical reasons, which enables English speakers 'to convey meanings more precisely and effectively for the right audience and context' (Liu and Espino, 2012), but also constitute a thorny area for EFL (English as a Foreign Language) learners because of their subtle nuances and variations in meaning and usage. Therefore, an important aspect of English linguistics is to distinguish one word from its synonyms or near-synonyms (Biber *et al.*, 1998; Divjak and Gries, 2006; Gries, 2001; Gries and Otani, 2010; Hanks, 1996; Liu, 2010; Xiao and McEnery, 2006).

This paper explores the collocational behavior and semantic prosodies of synonyms *in virtue of*, *owing to*, *thanks to*, *as a result of*, *due to* and *because of* in British National Corpus. While these words have a very similar denotational meaning, their collocational behavior and semantic prosodies differ markedly. The rest of this paper is structured as follows. In the next section, I will give an overview of related work by introducing corpus studies of collocation and semantic prosody. Section 3 introduces corpus data and tools used in this study. The results of this study are presented and analyzed in Section 4. The final section summarizes major findings and pedagogical implications of this study.

2. Related Work

Collocations are pervasive in texts of all genres and domains. Sinclair proposed that a collocation is a co-occurrence pattern that exists between two items that frequently occur in proximity to one another, but not necessarily adjacently or, indeed, in any fixed order. Node and collocates are two notions closely related to collocation. A node is an item whose total pattern of co-occurrence with other words is under examination; and a collocate is any one of the items which appears with the node within a specified span (Sinclair *et al.*, 2004).

The concept of semantic prosody was first introduced to the public by Louw (1993). From then on, it has become one of the important concepts in corpus linguistics and has claimed more and more attention from corpus linguists. 'Prosody' in the term 'semantic prosody' is borrowed from Firth (1957), who used it to refer to phonological coloring which spreads beyond semantic boundaries. Enlightened by Firthian sense of a 'prosody', Bill Louw coins the term 'semantic prosody' and endows it with its first definition, a 'consistent aura of meaning with which a form is imbued by its collocates' (Louw, 1993). Louw illustrates semantic prosody with the examples of *utterly*, *bent o*n and *symptomatic of,* which are followed by expressions referring to undesirable things (such as *destroying*, *ruining*, *clinical*, *depression*, *multitude of sins*, etc.). Michael Stubbs defines semantic prosody as 'a particular collocational phenomenon' and collocation as 'the habitual co-occurrence of two or more words' (Stubbs, 1996). Moreover, he classifies SP into three categorizations: negative prosody, positive prosody and neutral prosody. The node word bears a strong negative prosody if its collocates are mainly of strong negative semantic characteristics. The node word bears a strong positive prosody if its collocates are mainly of strong positive semantic characteristics. If both positive and negative collocates exist in the context, the node word can be said to bear a neutral or mixed prosody.

Following this categorization, Sinclair (1991) observed that *set in* has a negative prosody, typically collocating with words of unfavorable meanings, such as *rot, decay, malaise, ill-will, decadence, inflection, prejudice, vicious, rigour, mortis, numbness, bitterness, mannerism, anti-climax, anarchy*, and so on. Wei (2002a) found that *incidence* typically collocates with words referring to diseases or defects, such as *infections*, *HIV*, *injury*, *colon cancer*, *defects*, *disruption*, *cracking*, and so on. A few words have a positive prosody, for example, *provide* mainly collocates with words related to services or help, such as *facilities*, *information*, *services*, *aid*, *assistance*, *help*, *support*, *care*, *food*,

money, *nourishement*, *protection*, *security*, and so on Stubbs (1995). Many words are neutral or mixed in this respect. The main collocates of *probability of* can be divided into three categories: neutral collocates, such as *percent*, *change*, *occurrence* and *prototype structure*; positive collocates, such as *success*, *survival*, *acquisition*, and *winning*; negative collocates, such as *failure*, *accidents*, *error*, *loss*, *malfunction*, and *deficiency* (Wei, 2002b).

3. Method
3.1. Corpus Data: BNC
The British National Corpus (BNC) is a 100 million word collection of samples of written and spoken language from a wide range of sources, which is designed to represent a wide cross-section of British English from the later part of the 20th century, both spoken and written (Aston and Burnard, 1998). The written part of the BNC (90%) includes extracts from regional and national newspapers, specialist periodicals and journals for all ages and interests, academic books and popular fiction, published and unpublished letters and memoranda, school and university essays, among many other kinds of text. The spoken part (10%) consists of orthographic transcriptions of unscripted informal conversations and spoken language collected in different contexts, ranging from formal business or government meetings to radio shows and phone-ins.

BNC is monolingual, synchronic, general and sample-based by nature. It deals with modern British English, covers British English of the late twentieth century, includes many different styles and varieties instead of being limited to any particular subject field, genre or register, and that it contains many samples which allows for a wider coverage of texts within the 100 million limit. The corpus is encoded according to the Guidelines of the Text Encoding Initiative (TEI) to represent both the output from CLAWS (automatic part-of-speech tagger) and a variety of other structural properties of texts (e.g. headings, paragraphs, lists etc.). Full classification, contextual and bibliographic information is also included with each text in the form of a TEI-conformant header.

3.2. Corpus Tool and Analysis Procedure
The Sketch Engine (SkE) is a leading corpus tool, widely used in lexicography, language teaching, translation and the like (Kilgarriff *et al.*, 2004; Kilgarriff *et al.*, 2014). It includes two different things: the software, and the web service. The web service includes, as well as the core software, a large number of corpora pre-loaded and 'ready for use', and tools for creating, installing and managing users' own corpora. Corpora in SkE are often annotated with additional linguistic information, the most common being part of speech information (for example, whether something is a noun or a verb), which allows large-scale grammatical analyses to be carried out.

SkE has a number of core functions: Thesaurus, Wordlist, Concordance, Collocation, word sketches, and Sketch Diff. We are going to use Concordance, Collocation, word sketches and Sketch Diff functions in the present study. The span (the number of words left and right of the search word) is (-5, 5), the minimum frequency of each collocate being set 10 and minimum frequency in given range (in our case -5, 5) 5. Of seven measures to calculate the strength of collocation (T-score, MI, MI3, log likelihood, min. sensitivity, and logDice), I choose the default one *logDice* which is considered more reliable than the frequently used MI (mutual information) measure.

4. Results and Analysis
4.1. The Frequencies of in Virtue of Owing to Thanks to as A Result of Due to and Because of in Bnc

Table-1. Frequencies of *in virtue of*, *owing to*, *thanks to*, *as a result of*, *due to* and *because of* in BNC

	In Virtue of	Owing to	Thanks to	As A Result of	Due to	Because of
Total	110	681	1,766	4,468	10,189	15,922
Per million	0.98	6.06	15.72	39.77	90.69	141.70

As is shown in table 1, the frequency of *in virtue of* is the lowest, less than 1 in per million words. The frequencies of *owing to*, *thanks to*, *as a result of* and *due to* are much higher. The frequency of *because of* is the highest, more than 140 times of *in virtue of*.

4.2. The Collocates and Semantic Prosody of in virtue of in BNC

Table-2. The top 50 collocates of *in virtue of* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	5.45	19	9.98877	26	such	98560	0.40968
2	s.	2079	6.90394	27	right	74043	0.40675
3	solely	1637	6.22934	28	what	174892	0.39034
4	provisions	3998	5.99578	29	also	118806	0.14042
5	meaning	7784	5.51289	30	has	254360	0.04286
6	false	3410	4.8036	31	only	139969	-0.09588
7	object	5677	4.08636	32	two	140327	-0.09956
8	true	16385	3.99026	33	people	109954	-0.16302
9	rule	8324	3.54296	34	may	111435	-0.1823

10	legal	11609	3.48344	35	are	455086	-0.21117
11	rules	9616	3.33733	36	It	232134	-0.24032
12	4	28657	2.77284	37	not	432522	-0.40085
13	fact	36383	2.65202	38	(	403454	-0.45251
14	position	22521	2.53399	39	can	224417	-0.4546
15	common	17953	2.44421	40	or	360470	-0.45996
16	white	18401	2.40887	41	an	320093	-0.48127
17	being	86073	2.30532	42	do	242864	-0.56851
18	having	30419	2.10211	43	other	146437	-0.57604
19	something	47499	1.78298	44	this	343563	-0.58332
20	able	29603	1.72616	45	)	404196	-0.62509
21	things	39579	1.72355	46	;	204095	-0.63966
22	its	153975	1.46707	47	is	972921	-0.6442
23	us	60279	1.118	48	that	1062385	-0.7711
24	which	361203	1.06067	49	they	336549	-0.77597
25	their	245783	0.79277	50	have	453305	-0.79047

Table 2 lists the top 50 collocates of the phrase *in virtue of* in BNC automatically generated by the SkE. Further examination suggests that only 1 out of 50 (2%) collocates is positive: *true*. When *true* collocates with *in virtue of*, the meanings of the occurrences can be positive (1), neutral (2) or negative (3).

(1) One of the hypotheses which will be tested is that comprehension of the logical character of 'mental state verbs' is necessary for understanding that some sentences are true in virtue of their form alone (eg 'The father is a man').

(2) They are redundant, but not because the world can be described in terms of eternal" propositions which are true or false in virtue of being the propositions they are, but because they advertise certain claims which can equally successfully be conveyed implicitly, viz. by asserting the proposition or its negation, as the case may be.

(3) It is not true in virtue of any further relation between P and Q.

47 out of 50 (94%) collocates of *in virtue of* are neutral: *5.45*, *s.*, *solely*, *provisions*, *meaning*, *object*, *rule*, *legal*, *rules*, *4*, *fact*, *position*, *common*, *white*, *being*, *having*, *something*, *able*, *things*, *its*, *us*, *which*, *their*, *such*, *right*, *what*, *also*, *has*, *only*, *two*, *people*, *may*, *are*, *It*, *(*, *can*, *or*, *an*, *do*, *other*, *this*, *)*, *;*, *is*, *that*, *they* and *have*. When these words collocate with *in virtue of*, the meanings of a few occurrences are positive (4), and the meanings of most occurrences are neutral (5-8).

(4) He also thinks beauty is that in virtue of which we call beautiful things beautiful.

(5) To signifies this relation of subsequence in virtue of its potential meaning of a movement from one point in time to another and has been seen to give rise to two clearly identifiable actual meanings according to whether the speaker conceives the whole movement which to is capable of signifying or only the initial part thereof.

(6) It shows that something more must be said about what a convention is, about how much and what kind of agreement is necessary in order that a particular proposition of law can be true in virtue of a particular legal convention.

(7) The theory is that these audible or visible things have meaning only in virtue of expressing and evoking mental things, ideas or thoughts.

(8) Of course, if conventionalism were just the semantic theory that the phrase" legal rights" should not be used to describe rights people have in virtue of consistency in principle, then a conventionalist judge could indeed take a lively interest in that form of consistency under a different description.

2 out of 50 (4%) collocates of *in virtue of* are negative: *false* and *not*. When collocating with *in virtue of*, *false* only appears with *true*, stating some kind of condition, as in (9). The negative meaning of *false* is offset by the positive meaning of *true*, and the meaning of the sentence on the whole is neutral. The negative meaning of *not* renders the meaning of the sentence negative, as in (10).

(9) Analytic statements are those which relate concepts to concepts and are true (or false) solely in virtue of the meanings of their terms, for instance the statement that all bachelors are unmarried.

(10) Even if by chance we happened to conceive of just these members of the series, it is surely not in virtue of that fact that they count as correct continuations.

4.3. The Collocates and Semantic Prosody of Owing to in BNC

Table-3. The top 50 collocates of *owing to* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	probabilities	376	6.5392	26	cloud	1938	5.23016
2	lack	9543	6.53831	27	inadequate	2236	5.07469
3	sickness	1099	6.52427	28	difficulties	6638	4.96992
4	postponed	755	6.51216	29	absence	5631	4.96108
5	prevailing	905	6.36882	30	temporary	3628	4.92686
6	partly	5303	6.36	31	mainly	6888	4.92147

7	algae	550	6.31935	32	retired	2767	4.83342
8	wartime	982	6.30043	33	nature	15827	4.79648
9	cancelled	1178	6.13969	34	supplies	2974	4.74931
10	reductions	1210	6.11507	35	delay	3050	4.71962
11	fluctuations	753	6.09913	36	accident	5746	4.672
12	delays	1036	5.83929	37	fail	3263	4.63952
13	declining	1074	5.80771	38	increased	12655	4.61889
14	delayed	1668	5.80217	39	traffic	5997	4.61673
15	complexity	1711	5.776	40	wages	3455	4.57094
16	debts	1819	5.71229	41	introduction	4961	4.53801
17	negligence	1214	5.69698	42	debt	5068	4.5109
18	illness	3139	5.68561	43	funding	3991	4.39514
19	shortage	1449	5.52832	44	largely	7208	4.3763
20	weather	5406	5.42848	45	poor	12767	4.2849
21	restrictions	2518	5.35659	46	conclusion	4364	4.28432
22	necessity	1757	5.33348	47	escape	4405	4.27265
23	unable	5937	5.30782	48	falling	4435	4.26416
24	presence	7792	5.27327	49	injury	4515	4.24178
25	confusion	2714	5.2708	50	factors	8481	4.16048

As is shown in , 28 out of 50 (56%) collocates of *owing to* are neutral: *probabilities*, *prevailing*, *partly*, *algae*, *wartime*, *reductions*, *fluctuations*, *declining*, *complexity*, *weather*, *restrictions*, *necessity*, *presence*, *cloud*, *temporary*, *mainly*, *retired*, *nature*, *supplies*, *increased*, *traffic*, *wages*, *introduction*, *funding*, *largely*, *conclusion*, *falling* and *factors*. When these words collocate with *owing to*, the extended contexts render the meanings of the occurrences positive (11), neutral (12) or negative (13).

(11) She urged me to go with her to concerts and the theatre, and took me as her guest to the Edinburgh International Club which, owing to the presence of so many servicemen from overseas, was then an active and flourishing society.

(12) 'Generally, we bottle and sell them,' Apanage replied. 'Fole and Rattletrap go to the Crypticole Market - it is held infrequently and its location must of necessity change often, owing to its nature.

(13) These include possible reductions in care in the community owing to inadequate funding and the knock on effects on hospitals, which may face even greater difficulties in returning elderly people to the community than at present.

22 out of 50 (44%) collocates of *owing to* are negative: *lack*, *sickness*, *postponed*, *cancelled*, *delays*, *delayed*, *debts*, *negligence*, *illness*, *shortage*, *unable*, *confusion*, *inadequate*, *difficulties*, *absence*, *delay*, *accident*, *fail*, *debt*, *poor*, *escape* and *injury*. When these negative words collocate with *owing to*, the meanings of the occurrences are negative (14-17).

(14) With First National an optional insurance scheme is available to cover your monthly repayments if you are unable to work owing to accident, sickness or redundancy.

(15) The car was then marketed as 'Lagonda, designed under the supervision of W. O. Bentley', but the flood of orders could not be met owing to the postwar shortage of steel.

(16) The trouble is that there are just not enough places available at the moment, and in recent years local authorities have been unable to provide more owing to the shortage of funds and the cash limits imposed on local authority expenditure by central government.

(17) Land was reclaimed and settlements begun but then, owing to the poor quality of the ground following deforestation, they became the home of cattle.

4.4. The Collocates and Semantic Prosody of Thanks to In BNC

Table-4. The top 50 collocates of *thanks to* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	sincere	450	8.27858	26	1-0	328	5.96795
2	generosity	589	8.19081	27	ingenuity	338	5.96108
3	QUANTAS	15	8.10842	28	guests	3037	5.9402
4	Dougie	58	7.97446	29	unbeaten	502	5.8528
5	Australia's	349	7.86045	30	brighter	512	5.84645
6	airline	943	7.59646	31	Tea	515	5.84455
7	express	3804	7.20039	32	Sincere	11	5.78973
8	efforts	5363	7.10741	33	votes	3065	5.76189
9	Down	1784	7.01375	34	advances	1288	5.74544
10	Many	10850	6.86889	35	proceeding	681	5.7432
11	sponsorship	1094	6.84013	36	tremendous	1951	5.72504

12	Mirror	1040	6.71563	37	warmest	102	5.71768
13	heartfelt	144	6.68561	38	virtual	756	5.69965
14	grateful	2579	6.6153	39	vote	6883	5.62212
15	generous	2216	6.36265	40	contributed	2247	5.6145
16	extend	3039	6.35464	41	record	13862	5.59172
17	Under	4522	6.28851	42	Daily	2420	5.55361
18	Give	2008	6.28805	43	goals	4592	5.53557
19	largely	7208	6.26078	44	readers	4034	5.49816
20	Acknowledgements	30	6.18943	45	donors	413	5.49551
21	pioneering	544	6.14825	46	Our	9195	5.4868
22	sends	624	6.09913	47	surgeons	428	5.48561
23	owe	1201	6.05017	48	backing	1929	5.47057
24	goal	5783	6.02482	49	transplant	463	5.46278
25	high-tech	263	6.01345	50	Dr.	1215	5.45842

As is shown in table 4, 18 out of 50 (36%) collocates of *thanks to* are positive: *sincere*, *generosity*, *efforts*, *sponsorship*, *heartfelt*, *grateful*, *generous*, *Acknowledgements*, *pioneering*, *high-tech*, *ingenuity*, *unbeaten*, *brighter*, *Sincere*, *advances*, *warmest*, *donors* and *backing*. When these pleasant companies collocate with *thanks to*, the meanings of the occurrences are positive (18-21).

(18) My sincere thanks to you all for the most beautiful flowers that were sent to me in hospital during my recent operation.

(19) Lerwick Rugby Club players now sport a new strip, thanks to sponsorship from Honeywell, manufacturers of the TDC3000 distributed control system at Sullom Voe.

(20) HUNDREDS of thousands of dolphins could be saved from death thanks to pioneering work by British scientists.

(21) But today, Odessa's chances looked much brighter thanks to the intervention of a Northamptonshire headmaster who has brought hundreds of Bosnians to Britain.

32 out of 50 (64%) collocates of *thanks to* are neutral: *QUANTAS*, *Dougie*, *Australia's*, *airline*, *express*, *Down*, *Many*, *Mirror*, *extend*, *Under*, *Give*, *largely*, *sends*, *owe*, *goal*, *1-0*, *guests*, *Tea*, *votes*, *proceeding*, *tremendous*, *virtual*, *vote*, *contributed*, *record*, *Daily*, *goals*, *readers*, *Our*, *surgeons*, *transplant* and *Dr.* Although these words are neutral, the extended contexts render the meanings of the occurrences favorable, as is shown in (22) to (26).

(22) Now, we're looking for two couples to accompany us down under, to find out how you can perhaps qualify, stay tuned for the Dougie Down Under competition, thanks to QUANTAS, Australia's national airline.

(23) Part Two of Merlin's fabulous WWF Gold Series Collector cards hits the shops next week - and thanks to this super EXCLUSIVE Daily Mirror offer, you can get your hands on a whole pack of eight cards, worth 30p, absolutely free.

(24) By that date, largely thanks to his efforts, there were ten Welsh Jesuits, and fifty-seven priests in England and Wales.

(25) AMERICAN Independence Day at Bytes Restaurant, Honeywell, was a tremendous success thanks to three CCG youth trainees who were challenged by senior members of staff to organise a barbecue.

(26) It's thanks to the skill of **surgeons** and to the parents who allowed their childs heart to be donated.

4.5. The Collocates and Semantic Prosody Of As A Result of In BNC

Table-5. The top 50 collocates of *as a result of* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	suffered	5281	8.20839	26	Government's	2263	6.45332
2	partly	5303	7.87499	27	policies	8572	6.37374
3	largely	7208	7.84629	28	breach	3130	6.35652
4	accident	5746	7.68174	29	merger	1275	6.31977
5	suffer	3385	7.27043	30	lost	18253	6.28315
6	damage	8057	7.22037	31	arisen	708	6.24734
7	injuries	2441	7.10372	32	Partly	280	6.24633
8	loss	11122	7.04895	33	severe	4404	6.20688
9	changes	18220	7.00609	34	injured	3142	6.19378
10	incurred	1180	6.99591	35	changed	11059	6.16497
11	arise	3246	6.99363	36	pollution	3322	6.16006
12	occur	5539	6.84056	37	sustained	2002	6.1478
13	occurs	3278	6.83565	38	warming	914	6.0655
14	increased	12655	6.79433	39	increase	16661	6.05933
15	Gulf	2926	6.79172	40	brain	4383	6.0583
16	negligence	1214	6.77574	41	transactions	1955	6.05141
17	died	13558	6.72019	42	extinction	542	6.03133

18	crisis	5499	6.67125	43	decline	4599	6.02352
19	occurred	5361	6.54453	44	arising	2128	6.01306
20	losses	3696	6.5285	45	reduction	4690	6.00911
21	suffering	4132	6.51484	46	transfer	6239	6.00605
22	injury	4515	6.51087	47	reforms	2716	5.99678
23	pressure	11483	6.49807	48	factors	8481	5.9833
24	failure	7267	6.48147	49	arises	1774	5.97717
25	actions	4714	6.47926	50	suffers	742	5.97486

As is shown in table 5, 27 out of 50 (54%) collocates of *as a result of* are neutral: *partly*, *largely*, *changes*, *arise*, *occur*, *occurs*, *increased*, *Gulf*, *occurred*, *actions*, *Government's*, *policies*, *merger*, *arisen*, *Partly*, *changed*, *sustained*, *increase*, *brain*, *transactions*, *decline*, *arising*, *reduction*, *transfer*, *reforms*, *factors* and *arises*. When these words collocate with *as a result of*, the meanings of the occurrences can be positive (27), neutral (28) or negative (29).

(27) The improvements in the accident rate have come about largely as a result of increased awareness about safety matters and a developing safety culture.

(28) Assets are defined as 'rights or other access to future economic benefits controlled by an entity as a result of past transactions or events'.

(29) The Court of Appeal stated that it was impermissible to make a finding of 100% contributory negligence, as the Act states that the plaintiff must suffer damage partly as a result of his own fault and partly as a result of the defendant's fault.

23 out of 50 (46%) collocates of *as a result of* are negative: *suffered*, *accident*, *suffer*, *damage*, *injuries*, *loss*, *incurred*, *negligence*, *died*, *crisis*, *losses*, *suffering*, *injury*, *pressure*, *failure*, *breach*, *lost*, *severe*, *injured*, *pollution*, *warming*, *extinction* and *suffers*. When these negative words collocate with *as a result of*, the occurrences show unfavorable meanings, as is shown in (30) to (33).

(30) They [the dependants] are entitled to such a sum as will make good to them the financial loss which they have suffered and will suffer as a result of the death.

(31) The Act also stipulates that a manufacturer's guarantee cannot exclude liability for damage caused or loss suffered as a result of the manufacturer's negligence.

(32) Reports in December spoke of mounting concern over the effects of UN sanctions following government claims that up to 80,000 Iraqi children had died as a result of severe shortages in food and medicines.

(33) A study presented to the United Nations suggests that grain yields will fall by 10 to 15 per cent in much of the third world as a result of global warming.

4.6. The Collocates and Semantic Prosody of Due to in BNC

Table-6. The top 50 collocates of *due to* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	partly	5303	8.61992	26	changes	18220	6.80585
2	lack	9543	8.4933	27	injury	4515	6.79994
3	largely	7208	8.41329	28	This	110097	6.7496
4	mainly	6888	7.97366	29	next	41530	6.74247
5	failure	7267	7.609	30	poor	12767	6.74224
6	expire	230	7.53571	31	open	25357	6.72809
7	begin	7155	7.48771	32	completed	5872	6.70115
8	arrive	2800	7.42246	33	losses	3696	6.69819
9	fact	36383	7.35701	34	shortly	3126	6.69155
10	start	22299	7.35223	35	decline	4599	6.67142
11	increased	12655	7.23472	36	causes	4423	6.65699
12	loss	11122	7.15985	37	effects	10488	6.64154
13	error	3358	7.15696	38	demand	10686	6.62779
14	effect	22887	7.12605	39	problems	26466	6.57974
15	appear	10737	7.07326	40	charges	5932	6.57794
16	differences	7462	7.04228	41	before	79659	6.56161
17	increase	16661	7.04225	42	success	12805	6.55542
18	meet	13428	7.01162	43	accident	5746	6.53194
19	probably	24806	6.97123	44	month	13542	6.50991
20	factors	8481	6.96131	45	disease	8534	6.50792
21	damage	8057	6.95323	46	delay	3050	6.47742
22	illness	3139	6.94168	47	1993	5880	6.45427
23	retire	1064	6.82738	48	may	111435	6.43436
24	deaths	2322	6.81536	49	weather	5406	6.42876
25	entirely	6710	6.81333	50	difficulties	6638	6.42094

As is shown in table 6, 1 out of 50 (2%) collocates of *due to* is positive: *success*. *Success* collocating with *due to* typically shows a favorable evaluation, as in (34).

(34) The success of the project is due to the co-operation of local villagers, who collect up to 1,000 eggs a night.

31 out of 50 (62%) collocates of *due to* are neutral: *partly*, *largely*, *mainly*, *begin*, *arrive*, *fact*, *start*, *increased*, *effect*, *appear*, *differences*, *increase*, *meet*, *probably*, *factor*, *retire*, *entirely*, *changes*, *This*, *next*, *open*, *completed*, *shortly*, *causes*, *effects*, *demand*, *before*, *month*, *1993*, *may* and *weather*. When these words collocate with *due to*, the meanings of the occurrences can be positive (35), neutral (36) or negative (37).

(35) The trade balance on alcoholic drinks is in surplus, entirely due to a favourable trade balance in Scotch Whisky of £1.8 billion (Figure 5.5).

(36) The cooling effect is due to large amounts of material being injected into the stratosphere, preventing the sun's heat reaching the ground.

(37) The total loss due to these causes will depend not only on the method of construction and the geological formations over which the waterway passes, but also on the level of maintenance afforded.

18 out of 50 (36%) collocates of *due to* are negative: *lack*, *failure*, *expire*, *loss*, *error*, *damage*, *illness*, *deaths*, *injury*, *poor*, *losses*, *decline*, *problems*, *charges*, *accident*, *disease*, *delay* and *difficulties*. When these negative companies collocate with *due to*, the occurrences show negative meanings, as shown in (38) to (40).

(38) The problem of motivation may be due to the lack of self-esteem and self-actualisation the job gives the employee.

(39) Quite simply, this will ensure that the loan is repaid if you are unable to meet repayments due to death, illness or injury.

(40) The ADB blamed the slowdown on high inflation rates and interest rates, while the region had also been affected by reduced foreign investment, mainly due to Japan's financial difficulties.

4.7. The Collocates and Semantic Prosody of Because of in BNC

Table-7. The top 50 collocates of *because of* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	partly	5303	9.11838	26	increased	12655	6.75202
2	lack	9543	8.77316	27	pressure	11483	6.73358
3	nature	15827	7.92209	28	greater	13772	6.72708
4	its	153975	7.73422	29	illness	3139	6.7124
5	largely	7208	7.69235	30	lost	18253	6.62585
6	problems	26466	7.48661	31	uncertainty	2103	6.61718
7	mainly	6888	7.47682	32	importance	9456	6.59753
8	injury	4515	7.47552	33	weather	5406	6.57374
9	difficulties	6638	7.40889	34	but	321513	6.56944
10	their	245783	7.27854	35	financial	14204	6.56422
11	failure	7267	7.19929	36	suffered	5281	6.53491
12	poor	12767	7.18047	37	interest	26490	6.53471
13	difficulty	6170	7.13862	38	fear	8481	6.53375
14	size	12308	7.0977	39	commitments	1272	6.51534
15	high	32268	7.04347	40	cost	19937	6.51381
16	perhaps	24100	7.02438	41	precisely	3230	6.50271
17	risk	11465	7.00026	42	circumstances	10156	6.49938
18	fears	3106	6.97399	43	because	91030	6.49307
19	recession	3675	6.91158	44	economic	20117	6.41732
20	differences	7462	6.89616	45	costs	14775	6.41612
21	shortage	1449	6.89289	46	position	22521	6.41342
22	difficult	21447	6.81045	47	concern	9486	6.4114
23	simply	16796	6.80956	48	impact	7192	6.41041
24	low	15638	6.79566	49	probably	24806	6.40051
25	effects	10488	6.75415	50	involved	19627	6.39027

As is shown in table 7, 31 out of 50 (62%) collocates of *because of* are neutral: *partly*, *nature*, *its*, *largely*, *mainly*, *their*, *size*, *high*, *perhaps*, *differences*, *simply*, *low*, *effects*, *increased*, *greater*, *importance*, *weather*, *but*, *financial*, *interest*, *commiments*, *precisely*, *circumstances*, *because*, *econimic*, *costs*, *position*, *concern*, *impact*, *probably* and *involved*. When these words collocate with *because of*, the meanings of the occurrences can be positive (41), neutral (42) or negative (43).

(41) The water used in brewing Guinness comes from the catchment areas in the Wicklow Mountains close to where the Liffey rises and is particularly suited to the brewing of stout because of its purity and softness.

(42) I think that section of the world is the largest, and those countries contained within that section of the world have huge differences and that's because of religion, culture, different economies, different kinds of governments and so on and so forth.

(43) With legislative elections due in March or April, and presidential elections later in the year, Roh claimed that the postponement was necessary because of the country's current economic difficulties.

19 out of 50 (38%) collocates of *because of* are negative: *lack*, *problems*, *injury*, *difficulties*, *failure*, *poor*, *difficulty*, *risk*, *fears*, *recession*, *shortage*, *difficult*, *pressure*, *illness*, *lost*, *uncertainty*, *suffered*, *fear* and *cost*. When these unpleasant companies collocate with *because of*, the meanings of the occurrences are negative, as in (44) to (47).

(44) A plan to harness the hot springs of the Soufriere volcano on St Lucia, in the Caribbean, to generate geothermal energy has been abandoned because of unexpected technical problems and lack of funds.

(45) She can drink from a beaker with a lid and straw, but has to have someone with her because of the risk of choking.

(46) Others were not so fortunate, and the Director of the Laban Centre had to give up membership because of the pressure of time.

(47) Students in London have suffered particularly badly because of the Government's decision to make students ineligible for housing benefit.

4.8. Summary

Table-8. Distribution of *because of*-words across meaning categories in BNC

Synonyms	Positive	Neutral	Negative
in virtue of	0%	98%	2%
owing to	0%	56%	44%
thanks to	36%	64%	0%
as a result of	0%	54%	46%
due to	2%	62%	36%
because of	0%	62%	38%

As is shown in table 8, the semantic prosodies of these near synonyms vary. The semantic prosody of *in virtue of* is typically neutral. The semantic prosodies of *thanks to* is typically positive and neutral. The semantic prosodies of *owing to*, *as a result of*, *due to* and *because of* are typically neutral and negative. Besides, the collocations of these words also vary to a great extent.

5. Conclusion

This paper has explored the usage of *in virtue of*, *owing to*, *thanks to*, *as a result of*, *due to* and *because of* in British National Corpus with SkE. The results show that these near synonyms differ in their collocational behavior and semantic prosodies.

This study has a number of pedagogical implications. First of all, since near synonyms usually differ in their collocations and semantic prosodies, the traditional practice of explaining meanings to learners by offering synonyms should be used with caution. Teaching synonyms in this way can be a potential trap for learners which emphasizes the denotational meaning of words rather than their usage (Tognini-Bonelli, 2001). In practice, a DDL (data-driven learning) approach to learn or teach lexis in context should be proposed so that the learners could be aware of the collocations and semantic prosodies of a word. Then there will be fewer occurrences of unusual collocations and inappropriate semantic prosodies in their English. In addition, English dictionaries should incorporate information concerning the collocation and semantic prosody of a word or a phrase. Most contemporary learner dictionaries have not given enough attention to this aspect. The availability of large corpora and concordance software has made it possible to provide adequate information of typical collocations and semantic prosody for such dictionaries.

References

Aston, G. and Burnard, L. (1998). *BNC handbook: Exploring the British national corpus with sara.* Edinburgh University Press: Edinburgh.

Biber, D., Conrad, S. and Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use.* Cambridge University Press: Cambridge, UK.

Divjak, D. and Gries, S. T. (2006). Ways of trying in Russian: Clustering behavioral profiles. *Journal of Corpus Linguistics and Linguistic Theory,* 2(1): 23–60.

Firth, J. (1957). *Papers in linguistics.* Oxford University Press: Oxford.

Gries, S. T. (2001). A corpus-linguistic analysis of-ic and-ical adjectives. *ICAME Journal,* 25: 65–108. Available: http://www.hd.uib.no/icame/ij25/gries.pdf

Gries, S. T. and Otani, N. (2010). Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal,* 34: 121-50. Available: http://clu.uni.no/icame/ij34/gries_otani.pdf

Hanks, P. (1996). Contextual dependency and lexical sets. *International Journal of Corpus Linguistics,* 1(1): 75–98.

Kilgarriff, A., Rychly, P., Smrz, P. and Tugwell, D. (2004). *The sketch engine.* Proc Euralex, Lorient: France.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. and Suchomel, V. (2014). The sketch engine: Ten years on. *Lexicography,* 1(1): 7-36.

Liu, D. (2010). Is it a chief, main, major, primary, or principal concern? A corpus-based behavioral profile study of the near-synonyms. *International Journal of Corpus Linguistics,* 15(1): 56–87.

Liu, D. and Espino, M. (2012). Actually, Genuinely, Really, and Truly: A corpus-based Behavioral Profile study of near-synonymous adverbs. *International Journal of Corpus Linguistics,* 17(2): 198–228.

Louw, B. (1993). *Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies' in M. Baker, G. Francis, & E. Tognini-Bonelli (eds.): Text and Technology: In Honour of John Sinclair.* John Benjamins: Amsterdam. 157-76.

Sinclair, J. (1991). *Corpus, Concordance, Collocation.* Oxford University Press: Oxford.

Sinclair, J., Jones, S., Daley, R. and Krishnamurthy, R. (2004). English collocational studies: The OSTI Report. London: Continuum.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative methods. *Function of Language,* 2(1): 1-33.

Stubbs, M. (1996). *Text and corpus linguistics.* Blackwell: Oxford.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work.* John Benjamins: Amsterdam.

Wei, N. (2002a). Research methods in studies of semantic prosody. *Foreign Language Teaching and Research,* 34(2): 300-07.

Wei, N. (2002b). A corpus-driven study of semantic prosodies in specialized texts. *Modern Foreign Languages,* 25(1): 165-75.

Xiao, R. and McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A crosslinguistic perspective. *Applied Linguistics,* 27(1): 103-29.