Is It a Professional Career or Job? Using Sketch Engine to Investigate Synonyms

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

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

Correspondence: Bei Yang, School of English and Education, Guangdong University of Foreign Studies, Guangzhou 510420, China; TEL: 86-13535589695; E-mail: bei_yang@gdufs.edu.cn

ABSTRACT
As an important yet intricate linguistic feature in English language, synonymy poses a great challenge for second language learners. Using the 100 million-word British National Corpus (BNC) as data and the software Sketch Engine (SkE) as an analyzing tool, this article compares the usage of *career* and *job* by conducting the analysis of concordance, collocation, word sketches and sketch difference. The results show that different functions of SkE can make different contributions to the discrimination of *career* and *job*. The pedagogical implications of the findings are also discussed.

Keywords: career, job, BNC, Sketch Engine

INTRODUCTION

English is particular rich in synonyms due to historical reasons, which enables English speakers "to convey meanings more precisely and effectively for the right audience and context" (Liu & Espino 2012: 198). But synonyms also constitute a thorny area for EFL (English as a Foreign Language) learners because of their subtle nuances and variations in meaning and usage. Synonyms are not completely interchangeable. In fact, they differ in shades of meaning and vary in their collocations. Collocations are inaccessible to a speaker's conscious introspection (Hunston 2002: 142; Louw 1993: 173; Partington 1998: 68). However, with the development of corpora and corpora analysis tools, collocations have been addressed much more easily and frequently by linguists (Hunston 2002; Louw 1993, 2000; Partington 1998; Schmitt & Carter 2004; Sinclair 1991; Stubbs 1995, 1996, 2001; Xiao & Mcenery 2006).

The paper is structured as follows. Section two gives an overview of related work by introducing corpus studies of collocation, and its relevance to the study of synonyms. Section 3 introduces corpus data and tools used in this study. The results of this study are presented and analyzed in Section 4, where we show the success of Sketch Engine in researching synonyms. The final section summarizes major findings and pedagogical implications of this study.

RELATED WORK

Corpus studies of lexical semantics

The approach of using coprus evidence to study meaning of words or phrases is often labeled as corpus semantics or empirical semantics, and the most active and influential scholars are called neo-Firthian corpus linguists. The leading figure is John Sinclair who might as well be one of the first people to bring Firth's ideas together with a corpus linguistic methodology

(Stubbs 1996). Other important neo-Firthians include Michael Hoey, Susan Hunston, Bill Louw, Michael Stubbs, Wolfgang Teubert and Elena Tognini-Bonelli (McEnery & Hardie 2012: 122).

At the core of the neo-Firthian school of corpus linguistics is searching for the units of meaning. Inspired by Firth's (1957: 179) maxim that "you shall know a word by the company it keeps", Sinclair has paid much attention to the context in which a word is used. He firmly believes in the principle of 'trust the text' (Sinclair et al. 2004) and claims that 'the language looks rather different when you look at a lot of it at once' (Sinclair 1991: 100).

Reading concordance and calculating collocates from corpus are two important ways to study a lexical item in its context (Sinclair 1991). The concordance is the basic tool for anyone working with a corpus. Even far before the emergence of corpus linguistics, concordances to major works such as the Bible and Shakespeare have been available. With the help of computers, concordances are much easier to compile. For Sinclair (1991: 32), "A concordance is a collection of the occurrences of a word-form, each in its own textual environment. In its simplest form, it is an index. Each word-form is indexed, and a reference is given to the place of each occurrence in a text." In corpus linguistics, a simple and effective convention called KWIC (Key Word In Context) has been widely used.

Closely related to concordance is the notion of collocation. Collocation has been studied for at least five decades. Collocation was first used as a technical term by Firth (1957) when he said 'I propose to bring forward as a technical term, meaning by collocation, and apply the test of collocability' (Firth 1957: 194). According to Firth (1968: 181), 'collocations of a given word are statements of the habitual or customary places of that word.' Firth's research on collocation, however, is largely intuition-based. It is in sharp contrast with most corpus linguists' belief that the only way to reliably identify the collocates of a given word is to study patterns of co-occurrence in a corpus. For example, Hunston (2002, p. 68) argues, 'Collocation may be observed informally in any instance of language, but it is more reliable to measure it statistically, and for this a corpus is essential.' Sinclair operationalized the idea of Firth, proposing 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, p. 10).

Corpus-based studies of synonyms

In this paper, synonyms refer to lexical pairs that have very similar cognitive meanings or denotational meanings, but which may differ in collocations. Synonymous words, therefore, are not collocationally interchangeable (Tognini-Bonelli 2001: 34). For example, Halliday (1976: 73) observed that although *strong* and *powerful* share similar denotational meanings, tea is typically described as *strong* rather than *powerful* whereas a car is more likely to be described as *powerful* than *strong*. Gilquin (2003) investigates the difference between the English causative verbs *get* and *have*. Glynn (2007) compares intra- and extralinguistic factors in the contexts of *hassle*, *bother* and *annoy*. Gries & Otani (2010) study the synonyms *big*, *great* and *large* and their antonyms *little*, *small* and *tiny*. Other sets of synonyms that have attracted attention include *strong* and *powerful* (Church et al. 1991), *absolutely*, *completely* and *entirely* (Partington 1998), *big*, *large* and *great* (Biber et al. 1998), *quake* and *quiver* (Atkins & Levin 1995), *principal*, *primary*, *chief*, *main* and *major* (Liu 2010), and *actually*, *genuinely*, *really*, and *truly* (Liu & Espino, 2012)

METHOD

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, designed to represent a wide cross-section of British English from the later part of the 20th century, both spoken and written (Aston & Burnard 1998). The written part of the BNC (90%) includes, for example, 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, by nature, monolingual, synchronic, general and sample-based, in that it deals with modern British English, it covers British English of the late twentieth century, it 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.

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). It actually refers to 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.

RESULTS AND ANALYSIS

Synonyms of *Career*

In SkE the automatic identification of synonymy is achieved by the tool Thesaurus. SkE prepares a 'distributional thesaurus' for a corpus, a thesaurus created on the basis of common collocation. Two words will appear in each other's thesaurus entry if they have many collocates in common. For instance, if we find examples of both *professional career* and *professional job*, that is some evidence that the two nouns *career* and *job* are similar. These two words share the collocate *professional* in the modifier relation. In a very large computation, for all pairs of words, SkE computes how many collocates they share. The ones which share most are the ones that appear in a word's thesaurus entry. The thesaurus entry for the noun *career* is shown in Figure 1. From Figure 1, we can see that the word that share most collocates with *career* is *job*. In the following sections, we are going to make a comparison between the usage of *career* and *job*.

Figure 1. Synonyms of the noun *career*

The Frequencies of *Career* and *Job*

The concordance function of SkE enables researchers to compare frequencies of synonymous words. As shown in Table 1, the frequency of *job* is more than 3 times of *career*.

Table 1. Frequency of *career* and *job* in BNC

	career	job
total	8,955	31,262
Per million	79.71	278.27

The Collocates of *Career* and *Job*

Table 2. The top 50 collocates of *career* in BNC

Rank	Collocates	Freq.	logDice	Rank	Collocates	Freq.	logDice
1	pursue	4533	8.31403	26	political	29364	6.92397
2	distinguished	1101	8.19608	27	ruin	2707	6.91674
3	throughout	12249	7.82762	28	entire	4676	6.78896
4	successful	10680	7.75009	29	structure	18704	6.7837
5	academic	5359	7.62113	30	path	7154	6.77931
6	begin	40274	7.5877	31	guidance	2996	6.77708
7	prospect	5410	7.56477	32	promising	898	6.73365
8	his	410091	7.52108	33	graduate	2662	6.7055
9	opportunity	15556	7.51689	34	resume	1796	6.69528
10	professional	12550	7.51451	35	choice	13405	6.65989
11	embark	1335	7.50393	36	playing	1377	6.61936
12	officer	15222	7.35752	37	choose	17148	6.61347
13	stage	20995	7.3294	38	training	17548	6.60991
14	career	9101	7.21736	39	acting	571	6.5897
15	moral	5508	7.17992	40	racing	1780	6.56416
16	whose	19248	7.17064	41	artist	7483	6.54984
17	span	1333	7.16318	42	my	144750	6.53175
18	start	49642	7.12146	43	future	23328	6.51337
19	early	40920	7.11507	44	subsequent	4380	6.50652
20	progression	768	7.07474	45	education	19914	6.50469
21	launch	8471	7.04033	46	their	254439	6.45257
22	beginning	7490	7.03903	47	promotion	3499	6.44007
23	end	58747	7.01918	48	during	43497	6.42981
24	advancement	434	7.0127	49	football	5503	6.38987
25	development	32679	6.94915	50	spend	21412	6.37657

Table 2 shows the top 50 collocates of *career*. The dominant collocates of *career* can be grouped into four grammatical categories: nouns, verbs, adjectives and function words. 23 out of 50 (46%) collocates of *career* are nouns: *prospect, opportunity, officer, stage, career, progression, beginning, end, advancement, development, structure, path, guidance, graduate, choice, playing, training, acting, racing, artist, education, promotion* and *football*. Of the above collocates, the meanings of some nouns are positive, such as *prospect, opportunity, progression, advancement, development, guidance* and *promotion*. When these pleasant companies collocate with *career*, the meanings of the occurrences are positive, as in (1) to (3). The meanings of the rest nouns are neutral, which render the meanings of the occurrences neutral too, as in (4) to (6).

(1) I commend her to you without reservation - she is an outstanding professional with excellent career **prospects**, and will be an asset to any library authority.

(2) In Britain also, several types of paraprofessional training programmes have been developed that provide useful avenues for career **advancement**.

(3) It is the policy of the Group to afford disabled persons full and fair consideration for employment and subsequent training, career development and **promotion** on the basis of their aptitudes and abilities.

(4) 'I used to know Wapping well at one **stage** of my career,' Devlin said.

(5) A typical career **path** might be a young European starting in the hotel industry as a management trainee, gaining experience in a variety of establishments in the far East and Europe, then progressing from, say, a major hotel chain through the food and beverage side to a management position at the top end of the leisure market.

(6) When a **football** career ends and reality tackles back, many footballers are attracted like a magnet to the world they know best.

10 out of 50 (20%) collocates of *career* are verbs: *pursue, begin, embark, span, start, launch, ruin, resume, choose* and *spend*. When *pursue, begin, embark, span, start, launch* and *resume* collocate with *career*, the extended contexts render the meanings of the occurrences favorable, as in (7) to (9). When *ruin* collocates with *career*, however, the meanings of the occurrences are negative, as in (10). *Choose* and *spend* are neutral, and when they collocate with *career* the meanings of the occurrences are neutral, as in (11).

(7) As a center of philosophical activity, Edinburgh remains at the forefront in the British Isles, and many of its postgraduates have gone on to **pursue** academic careers around the world.

(8) Tony Rudd retires this month after a distinguished career **spanning** 53 years - 13 at Rolls-Royce, 18 at BRM and 22 at Lotus.

(9) The two had a son, Neville, and it was not until after the war years that she was in a position to **resume** her golfing career.

(10) He tells of one teacher who was interviewed by police for 12 hours, and then suspended from work for six months after one of his pupils accused him of touching her up. Eventually the charges were dropped, but the teacher's career was **ruined**.

(11) The answers to these questions will have important consequences for anyone who is about to **choose** a career or a potential employer.

11 out of 50 (22%) collocates of *career* are adjectives: *distinguished, successful, academic, professional, moral, early, political, entire, promising, future* and *subsequent*. Some adjectives are positive, such as *distinguished, successful, professional* and *promising*. When these positive words collocate with *career*, the meanings of the occurrences are pleasant, as in (12) and (13). The rest adjectives are neutral, which render the meanings of the occurrences neutral, as in (14) and (15).

(12) JOAN CROSS can be said to have had two **distinguished** performing careers: pre-war

and post-war.
13) Building on the TV success that rocketed her to overnight fame, she is set to make TWO movies and is launching a **promising** pop career.
14) I used to think about an **academic** career when I was a student. That was before I got married, of course.
15) Cubby (seated centre) spent his **entire** career in Dundee, starting as an apprentice at Westport Branch.

The remaining 6 (12%) collocates of *career* are function words: *throughout*, *during*, *his*, *whose*, *my* and *their*. When these words are used with career, the meanings of the occurrences are neutral, as in (16).

(16) Crawford explained, '**Throughout** my acting career , I have always taken notes of myself from directors and from actors.

From the above analysis, we can see out of the 50 collocates of *career*, 36% are positive, 62% are neutral and 2% are negative.

Table 3. The top 50 collocates of *job* in BNC

Rank	Collocates	Freq.	logDice	Rank	Collocates	Freq.	logDice
1	loss	15201	9.03148	26	opportunity	15556	7.21993
2	lose	28242	8.95948	27	thousand	15369	7.19144
3	create	21419	8.68787	28	another	59114	7.17297
4	job	31353	8.57753	29	for	880813	7.16771
5	get	208360	8.24464	30	industry	21212	7.15357
6	do	534989	8.14896	31	a	2163730	7.15338
7	good	137820	8.12248	32	whose	19248	7.1439
8	description	6455	8.00653	33	give	129248	7.11709
9	offer	35352	7.84374	34	pay	42524	7.07211
10	new	105682	7.67824	35	security	12069	7.06449
11	satisfaction	2842	7.65659	36	properly	5523	7.04262
12	their	254439	7.55709	37	because	100637	7.04077
13	creation	4780	7.53503	38	interview	9084	7.03078
14	cut	24247	7.53419	39	keep	49239	6.9979
15	part-time	2083	7.49838	40	more	209635	6.96429
16	worker	17220	7.49553	41	would	278265	6.94269
17	my	144750	7.47253	42	have	1302008	6.91585
18	find	96211	7.43249	43	his	410091	6.91109
19	your	133466	7.42516	44	work	138402	6.9038
20	training	17548	7.34397	45	manager	17060	6.89904
21	apply	19661	7.32397	46	hundred	22353	6.88956
22	people	121326	7.26856	47	finish	13351	6.88574
23	full-time	2169	7.24723	48	part	64918	6.87282
24	temporary	3735	7.24491	49	will	330553	6.86938
25	take	174006	7.23702	50	want	88159	6.84855

As is shown in Table 3, the dominant collocates of *job* can be grouped into four grammatical categories: nouns, verbs, adjectives and function words. 16 out of 50 (32%) collocates of *job* are nouns: *loss*, *job*, *description*, *satisfaction*, *creation*, *worker*, *training*, *people*, *opportunity*, *thousand*, *industry*, *security*, *work*, *manager*, *hundred* and *part*. Of the above nouns, several words are positive, such as *satisfaction*, *creation*, *opportunity* and *security*. These positive collocates render the meanings of the occurrences pleasant, as in (17) and (18). Some words are neutral, such as *job*, *description*, *worker*, *training*, *people*, *thousand*, *industry*, *work*, *manager*, *hundred* and *part*. When these neutral collocates occur with *job*, the meanings of the

occurrences are also neutral, as in (19) and (20). The unpleasant meaning of *loss*, however, makes the meanings of the occurrences negative, as in (21).

(17) The final system design is evaluated on the basis of job **satisfaction** of those working on it as well as its efficiency.

(18) Graduates enjoy greater job **opportunities** than those entering employment direct from school.

(19) Erm as to the assertion that Harrogate wants office jobs and not industrial **jobs**, and I think the main point there is that we we're simply trying to achieve jobs to meet the needs of our resident workforce.

(20) We need more money to improve transport in London, and provide jobs where **people** need them.

(21) However White and Mackay has ruled out any major job **losses** from the takeover.

15 out of 50 (30%) collocates of *job* are verbs: *lose*, *create*, *get*, *do*, *offer*, *cut*, *find*, *apply*, *take*, *give*, *pay*, *interview*, *keep*, *finish* and *want*. Of the above verbs, the meanings of some words are neutral (*create*, *get*, *offer*, *find*, *take*, *give* and *keep*), but the extended contexts render the meanings of the occurrences positive, as in (22) and (23). The meanings of *lose* and *cut*, on the other hand, are negative, and when these two words occur with *job*, the meanings of the occurrences are unfavorable, as in (24).

(22) Speaking at the launch, party leader John Smith said that adapting to environmental changes could **create** 700,000 jobs and generate business worth £140,000 million by the end of the century.

(23) She **found** her father a job - as the internal postman.

(24) The point that this motion makes is to try to make action, the facilitating attitudes over there is gonna lead to five hundred people **losing** their jobs in April.

8 out of 50 (16%) collocates of *job* are adjectives: *good*, *new*, *part-time*, *full-time*, *temporary*, *another*, *properly* and *more*. The meanings of *good* and *properly* are pleasant, and when these words collocate with *job*, the meanings of the occurrences are positive, as in (25). The meanings of the other adjectives are neutral, and the collocation of these words with *job* render the meanings of the occurrences neutral (26).

(25) The President has done a **good** job of putting himself in the position where Arctic drilling is seen as patriotic", she said.

(26) Grace was juggling a **full-time** job with looking after three children when she got the idea to set up her business.

11 out of 50 (22%) collocates of *job* are function words: *their*, *my*, *your*, *for*, *a*, *whose*, *because*, *would*, *have*, *his* and *will*. These function words are all neutral. When these function words collocate with *job*, the meanings of the occurrences are neutral (27).

(27) I think most of the issues that have been raised tonight already, are ones for the police authority and I hope they will get on with **their** job .

From the above analysis, we can see that of the 50 collocates of *job*, 26% are positive, 68% are neutral and 6% are negative.

The Syntactic Patterns of *Career* and *Job*
The function that gives the Sketch Engine its name is the word sketch: a one-page summary of a word's grammatical and collocational behavior. Figure 2 demonstrates part of the word sketch for *career*. Its collocates are grouped according to grammatical relations in which they occur. For example, in the first column, a number of words such as *distinguished*, *successful*,

academic, *playing* and *moral* are grouped under the category of modifiers of "career". Figure 3 shows part of the word sketch for *job*.

Figure 2. Word sketch for the noun *career* in BNC

Figure 3. Word sketch for the noun *job* in BNC

In order to present a fine-grained comparison, we summarized the 10 patterns of *career* and 15 patterns of *job* in table 4 and table 5.

Table 4. The syntactic behavior of *career* in BNC

| Categories	Score	Example
modifiers of "career"	35.48	has had a **distinguished career** at the British Museum
nouns and verbs modified by "career"	19.02	for those seeking skills for further **career advancement**
verbs with "career" as object	25.90	He **began** his **career** originally as a journalist
verbs with "career" as subject	13.43	Norman Schwarzkopf's army **career spanned** 35 years
"career" and/or…	8.21	You expected me to drop my **life** and my **career** for you
prepositional phrases	*	In the early stages **of** his political **career**
adjective predicates of "career"	1.14	The notion of a **career** freely **open** to all talents was
possessors of "career"	7.56	A major retrospective of the **artist's career** is scheduled to
verbs with particle "up" and "career" as object	0.91	Cindy Gallop has given **up** her **career** in theatre marketing
verbs with particle "out" and "career" as object	0.22	Dennis carved **out** a **career** in the building industry

Table 5. The syntactic behavior of *job* in BNC

Categories	Score	Example
modifiers of "job"	31.98	The President has done a **good job** of putting himself
nouns and verbs modified by "job"	13.09	there will be further **job losses** - around 250
verbs with "job" as object	40.93	how the clerks **did** their **job**
verbs with "job" as subject	12.04	My **job entailed** being on call for shipping
"job" and/or…	7.42	French people would take priority for **jobs and housing**
prepositional phrases	*	She does a variety **of jobs** for learned bodies in law
adjective predicates of "job"	1.62	the new system will make their **job** less **secure**
"job" is a…	0.45	ending of temporary **jobs was the reason** for unemployment
possessors of "job"	2.30	he was offered the **manager's job** at Birmingham City
usage patterns	0.27	**it's** your **job to** take immediate steps
Verbs with particle "up" and "job" as object	0.70	What in fact he did was give **up** his regular **job**
…is a "job"	0.57	The librarian's **job** is a **job** of management of information
verbs with particle "out" and "job" as object	0.18	they still are unable to carry **out** the **job**
verbs with particle "down" and "job" as object	0.15	She managed to hold **down** a **job** as a journalist
verbs with particle "over" and "job" as object	0.03	waiting for a robot to take **over** your **job**

It has to be noted that although *career* and *job* share many similarities in their syntactic patterns, there are also apparent differences, as is shown by the Sketch Diff function of SkE.

Comparison of Lexical and Grammatical Collocates of *Career* and *Job*
The Sketch Diff function of SkE allows users to visually compare synonymous words based on their grammatical relations and collocational behaviors. Figure 4 shows part of the differences between *career* and *job* automatically generated by SkE. In the figure, the greener a word is, the more closely it relates to *career*. The redder a word is, the more closely it relates to *job*. For example, in the "career/job" and/or pattern, *reputation*, *background*, *relation* and *interest* frequently collocate with *career*, but are never used with *job*. On the other hand, *income*, *status*, *growth* and *profession* always collocate with *job*, but are never used with *career*. There are some other words that can occur both with *career* and *job* in the "career/job" and/or relation, such as *life*, *marriage*, *relationship*, *employment*, *family*, *career*, *education*, *money*, *training*, *job* and *home*.
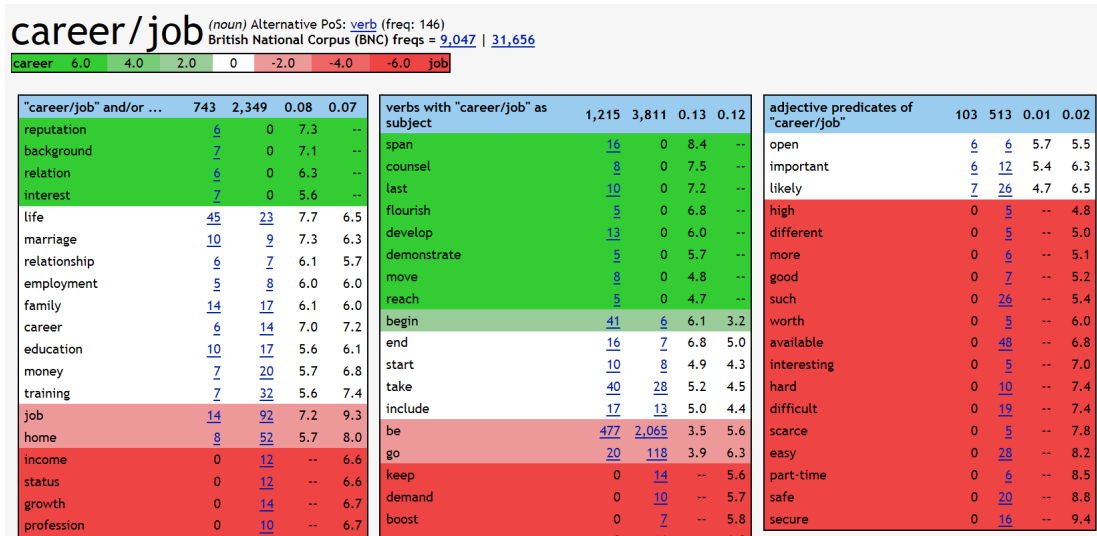
In order to make the comparison results clearer, we present part of the differences in the following tables. As is shown in table 6, only *career* can be the subject of some verbs, such as *span, counsel, last, flourish, develop, demonstrate, move* and *reach*, as in (28). Only *job* can be subject of some verbs, such as *keep, demand, boost, create, axe, lose, disappear, require, involve* and *entail*, as in (29). Both *career* and *job* can be the subject of some verbs, such as *begin, end, start, take, include, be* and *go*. Take '*go*' as an example, it collocates with *career* for 20 times, and collocates with *job* for 118 times, as in (30) and (31).

Table 6. verbs with "career/job" as subject

	frequency	frequency	score	score
span	16	0	8.4	--
counsel	8	0	7.5	--
last	10	0	7.2	--
flourish	5	0	6.8	--
develop	13	0	6.0	--
demonstrate	5	0	5.7	--
move	8	0	4.8	--
reach	5	0	4.7	--
begin	41	6	6.1	3.2
end	16	7	6.8	5.0
start	10	8	4.9	4.3
take	40	28	5.2	4.5
include	17	13	5.0	4.4
be	477	2,065	3.5	5.6
go	20	118	3.9	6.3
keep	0	14	--	5.6
demand	0	10	--	5.7
boost	0	7	--	5.8
create	0	16	--	6.0
axe	0	9	--	6.2
lose	0	23	--	6.3
disappear	0	16	--	6.5
require	0	39	--	6.9
involve	0	42	--	7.0
entail	0	20	--	7.3

(28) Alan Crosskill, general service manager and Press spokesman for Cleveland ambulance, is leaving the service after ten years to become self employed in public relations and **career counselling**.

(29) The only possibilities for such areas might lie in attracting to them (through government policy) low-paid **jobs demanding** minimal skills.

(30) Justin hasn't had this much ink for a decade, not since his million-pound transfer from Norwich to Nottingham Forest, where, from being the highest scorer in the First Division, his **career went** into a premature tailspin.

(31) Over the years 's **job** has **gone** from unloading individual bags to today's bulk deliveries of around 40,000 tonnes each year.

As is shown in table 7, some adjectives only modify *career*, such as *distinguished, successful, playing, acting, moral, promising, racing, entire, subsequent* and *test*, as in (32). Some adjectives only modify *job*, such as *excellent, proper, odd, manufacturing, difficult, permanent, top, temporary* and *part-time*, as in (33). Some adjectives modify *career* and *job*, such as *academic, brilliant, professional, full-time, new* and *good*. Take '*professional*' as an example, it collocates with *career* for 67 times, and it collocates with *job* for 52 times, as in (34) and (35).

Table 7. modifiers of *career/job*

	frequency	frequency	score	score
distinguished	82	0	9.3	--
successful	91	0	8.3	--
playing	29	0	8.1	--
acting	28	0	8.1	--
moral	61	0	8.1	--
promising	25	0	7.7	--
racing	24	0	7.6	--
entire	41	0	7.5	--
subsequent	33	0	7.2	--
test	31	0	7.2	--
academic	64	9	8.2	4.4
brilliant	24	9	7.2	4.6
professional	67	52	7.6	6.5
full-time	9	119	5.9	8.4
new	94	582	5.0	7.5
good	22	945	3.2	8.5
excellent	0	58	--	7.0
proper	0	74	--	7.3
odd	0	64	--	7.4
manufacturing	0	66	--	7.4
difficult	0	78	--	7.5
permanent	0	77	--	7.5
top	0	109	--	7.5
temporary	0	106	--	8.0
part-time	0	122	--	8.4

(32) May I begin by paying tribute to the long and **distinguished career** of the late Baroness Hart of South Lanark? No one in the developing world or the ODA will forget her great contribution.

(33) He had a **part-time job** in a large firm of chemists where he worked in the stockrooms. There had been rumours of redundancies.

(34) They had presumably been drawn to Bologna as students and had stayed there to carve themselves out a teaching or **professional career**.

(35) Meanwhile 58.3 per cent of the sons of non-skilled manual workers stayed working class, with only 10.7 per cent moving up into managerial or **professional jobs**.

As is shown in table 8, in the *career/job in* category, some words only collocate with *career*, such as *journalism, medicine, engineering, advertising, science, nursing, football, navy, management, politics, commerce* and *theatre,* as in (36). Some words only collocate with *job*, such as *government, store, economy, jeopardy, shop, manufacturing, office, hour, hand, sector* and *factory*, as in (37). Some words not only collocate with *career* but also with *job*, such as *service* and *industry*. Take '*service*' as an example, it collocates with *career* for 20 times, and also collocates with *job* for 20 times, as in (38) and (39).

Table 8. career/job in

	frequency	frequency	score	score
journalism	13	0	9.2	--
medicine	11	0	8.6	--
engineering	9	0	8.4	--
advertising	8	0	8.3	--
science	15	0	8.3	--
nursing	7	0	8.3	--
football	8	0	8.2	--
navy	7	0	8.1	--
management	12	0	8.1	--
politics	12	0	7.9	--
commerce	5	0	7.8	--
theatre	7	0	7.7	--
service	20	20	8.1	7.7
industry	25	77	7.9	9.3
government	0	10	--	7.0
store	0	8	--	7.1
economy	0	11	--	7.2
jeopardy	0	7	--	7.3
shop	0	15	--	7.5
manufacturing	0	9	--	7.5
office	0	22	--	7.5
hour	0	15	--	7.7
hand	0	51	--	8.0
sector	0	42	--	8.8
factory	0	27	--	8.9

(36) Once out of the Army, John Moynihan, after a couple of false starts, began his **career in journalism**, before long landing a job with the Evening Standard on their 'In London Last Night' column.

(37) In 1940 I was fortunate to find a summer **job in** the local tomato **factory**, one of the innumerable similar factories scattered in the Parma province.

(38) Sheila Mossman, who as part of the lifetime **career in** the **service** of music was an examiner with the board, died in 1971.

(39) Again it has been shown that **jobs in services**, clerical work, agriculture and construction are attractive, but factory work is only the fifth choice, preferred by only 11 per cent (see figure 3.7).

Limitations of SkE

So far we have demonstrated how to use some core functions of SkE to research synonymous nouns *career* and *job*. However, it has to be pointed out that SkE has not without its limitations. One apparent limitation is its automatic extraction of similar words. In Figure 1, some of the synonyms provided by the tool Thesaurus seems to have little similarities with *career*, such as *education*, *success*, *history*, *season*, *game*, etc. A recent study carried out by Perisman et al. (2015) on how to automatically identifying and extracting synonyms might be able to help SkE to improve its accuracy. In addition, SkE cannot semantically annotate a corpus as another web-based corpus tool Wmatrix does. The SkE team may wish to solve this problem in the future.

CONCLUSION

Researching synonymy is a crucial task in the field of lexical semantics because of its importance and intricacy. In this paper, we have introduced the leading corpus tool SkE and its

advantages in investigating synonymous verbs. The results show that different functions of SkE can make different contributions to the discrimination of *career* and *job*.

This study has a number of pedagogical implications. First, from the above analysis we can see that synonyms usually differ a lot in their collocations, so the traditional practice of explaining meanings to learners by offering synonyms should be used very carefully. 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: 34). Second, second language acquisition studies show that native-speakers memorize lots of chunks of words, and these ready-made or prefabricated units contribute to naturalness and fluency of their utterances. Therefore, if EFL learners want to achieve native-like selection and native-like fluency, they also need to store the collocational patterns from Table 2 to 3. Third, given the huge number of synonyms in English, teachers cannot teach the collocational behaviour and semantic prosody of all the synonyms to students. Teach students how to use SkE to conduct their own research is a better solution.

References

Atkins, B., & Levin, B. (1995). Building on a corpus: A linguistic and lexicographical look at some near-synonyms. *International Journal of Lexicography, 8*(2), 85–114.

Aston, G., & Burnard, L. (1998). *The BNC Handbook: Exploring the British National Corpus with SARA*. Edinburgh University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge, UK: Cambridge University Press.

Church, K. W., Gale, W., Hanks, P., & Hindle, D. (1991). Using statistics in lexical analysis. In U. Zernik (Ed.), *Lexical Acquisition: Exploiting On-line Resources to Build a Lexicon* (pp. 115–164). Hillsdale, NJ: Lawrence Erlbaum.

Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. In Philological Society (Eds.), *Studies in Linguistic Analysis* (pp. 1–32). Oxford, UK: Blackwell.

Firth, J. R. (1968). A Synopsis of Linguistic Theory. In F.R. Palmer (Ed.), *Selected Papers of J.R. Firth 1952–59* (168–205). London: Longmans.

Gilquin, G. (2003). Causative 'get' and 'have': So close, so different. *Journal of English Linguistics, 31*(2), 125–148.

Glynn, D. (2007). *Mapping meaning. Towards a usage-based methodology in Cognitive Semantics*. (Unpublished doctoral dissertation). University of Leuven, Leuven, Belgium.

Gries, S. Th., & Otani, N. (2010). Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal*, *34*, 121–150.

Halliday, M. A. K. (1976). *System and Function in Language*. Oxford: Oxford University Press.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). *The Sketch Engine*. Proc Euralex, Lorient, France.

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., & 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*. Amsterdam: John Benjamins, pp. 157-76.

Louw, B. (2000). 'Contextual prosodic theory: Bringing semantic prosodies to life' in C. Heffer, H. Sauntson, and G. Fox (eds.): *Words in Context: A Tribute to John Sinclair on his Retirement*. Birmingham: University of Birmingham.

McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

Partington, A. (1998). *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam, Netherlands: John Benjamins.

Schmitt, N., & Carter, R. (2004). 'Formulaic sequences in action: An introduction' in N. Schmitt (ed.): *Formulaic Sequences*. Amsterdam: John Benjamins, pp. 1-22.

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

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, 1–33.

Stubbs, M. (1996). *Text and Corpus Linguistics*. Oxford: Blackwell.

Stubbs, M. (2001). *Words and Phrases*. Oxford: Blackwell.

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

Xiao, R., & T, McEnery. (2006). Collocation, semantic prosody, and near synonymy: A crosslinguistic perspective. *Applied Linguistics*, 27(1), 103-129.