English Language Teaching; Vol. 9, No. 1; 2016
ISSN 1916-4742 E-ISSN 1916-4750
Published by Canadian Center of Science and Education

A Corpus-based Comparative Study of *Learn* and *Acquire*

Bei Yang[1]

[1] School of English and Education, Guangdong University of Foreign Studies, Guangzhou, China

Correspondence: Bei Yang, Associate Professor, School of English and Education, Guangdong University of Foreign Studies, No. 2 Baiyun Avenue (North), Baiyun District, Guangzhou 510420, China. Tel: 86-20-3932-8080. E-mail: bei_yang@gdufs.edu.cn

Received: November 12, 2015 Accepted: December 20, 2015 Online Published: December 21, 2015

doi:10.5539/elt.v9n1p209 URL: http://dx.doi.org/10.5539/elt.v9n1p209

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 *learn* and *acquire* used in natural discourse 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 *learn* and *acquire*. Pedagogical implications are discussed when the results are introduced into the classroom.

Keywords: learn, acquire, BNC, Sketch Engine

1. Introduction

One day, a student asked me: "The verbs *learn* and *acquire* share the following similar meaning: to develop or gain knowledge and skill. Why do we say *acquire* knowledge instead of *learn* knowledge, and why do we say *learn* to drive instead of *acquire* to drive?" I answered: "*Learn* and *acquire* are synonyms. They share similar meanings and usages, but they also differ in collocational and colligational patterns." The student continued to ask: "What are the collocational and colligational patterns of *learn* and *acquire* respectively?" Being a second language learner myself, I found it hard to give her a satisfactory answer. Therefore, I went to the library and tried to find the answer from reference books. In *Merriam Webster's Dictionary of Synonyms*, *learn* and *acquire* are not classified as synonyms. *Longman Synonym Dictionary* lists the synonyms of *learn* and *acquire* respectively, but offers no further explanation as to the similarities and differences between the two verbs. Unable to find the satisfactory answer, I decided to conduct a corpus-based comparative study of *learn* and *acquire* to address the perplexing question.

The paper is structured as follows. Section two gives an overview of related work by introducing corpus studies of collocation and colligation, and their 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 I show the success of Sketch Engine in researching synonyms. The final section summarizes major findings and pedagogical implications of this study.

2. Related Work

2.1 Corpus Studies of Collocation and Colligation

Collocations are pervasive in texts of all genres and domains. Although the study of collocations can be traced back to ancient Greece, the notion of collocations was first brought up by Palmer (1933) in English language teaching and later introduced to the field of theoretical linguistics by Firth (1957). The often-cited definition of collocations is "statements of the habitual and customary places of that word" (Firth, 1957, p. 181).

Nevertheless, Firth's research on collocation is largely intuition-based, which 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." The idea that Firth proposed is operationalized by Sinclair and associates' early work from 1970s and later collocation becomes one of a few most important concepts in corpus linguistics. Collocations impose a great challenge for second language learners. Numerous studies indicate that learners' language is problematic in the

idiomatic usage of English, which can be mainly attributed to misrepresented collocations.

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. Closely related to collocation is the notion of node and collocates. 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). Collocates are also determined within particular spans: "Two other terms ... are span and span position. In order that these may be defined, imagine that there exists a text with types A and B contained in it. Now, treating A as the node, suppose B occurs as the next token after A somewhere in the text. Then we call B a collocate at span position +1. If it occurs as the next but one token after A, it is a collocate at span position +2, and so on." (Sinclair et al., 2004, p. 34)

In order to test whether two words are significant collocates, four pieces of data are required: the length of the text in which the words appear, the number of times they both appear in the text, and the number of times they occur together (Sinclair et al., 2004, p. 28). Building on Sinclair's work, Hoey (2005, p. 5) defines collocation as "a psychological association between words (rather than lemmas) up to four words apart and is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution".

The notion of colligation is closely related to that of collocation. The term colligation was introduced by Firth (1968, p. 181) in order to distinguish lexical interrelations from those holding between grammatical categories:

The statement of meaning at the grammatical level is in terms of word and sentence classes or of similar categories and of the interrelation of those categories in colligations. Grammatical relations should not be regarded as relations between words as such – between *watched* and *him* in 'I watched him' – but between a personal pronoun, first person singular nominative, the past tense.

Hoey provides a straightforward definition: colligation can be defined as 'the grammatical company a word keeps and the positions it prefers'; in other words, a word's colligations describe what it typically does grammatically' (Hoey, 2005). Thus, colligation is a similar idea to collocation, but with a different emphasis. For example, 'verb + *to* infinitive' is a colligation, while *dread* + *think* is a collocation which exemplifies the colligation. Irrespective of the definition adopted, colligation, like collocation, is a probabilistic relation.

2.2 Corpus Approaches to Synonyms

Synonymy, or semantic equivalence, is an important yet intricate linguistic feature in the field of lexical semantics. Synonyms are not completely interchangeable; rather, they differ in shades of meaning and vary in their connotations, implications, and register (DiMarco et al., 1993). Any natural language consists of a considerable number of synonymous words. 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, p. 198), but also constitute a thorny area for EFL (English as Foreign Language) learners because of their subtle nuances and variations in meaning and usage.

It thus comes no surprise that an important aspect of English linguistics is to find the proper measures of automatically identifying and extracting synonyms (Peirsman, Geeraerts & Speelman, 2015) and of distinguishing one word from its synonyms or near-synonyms (Hanks, 1996; Biber et al., 1998; Gries, 2001; Xiao & McEnery, 2006; Divjak, 2006; Gries & Otani, 2010; Liu, 2010; Hu & Yang, 2015). Although the two orientations of researching synonyms are equally important, I will in this paper focus more attention on the second one. I would like to discover what the relative strengths and weaknesses of using Sketch Engine to research synonyms are, and what their relative scope of applicability is.

The past decades have witnessed significant advances in the studies on synonymy, which was boosted by the advent of the computer era and the central ideas of corpus semantics. Based on the Brown Corpus, Miller & Charles (1991) find that the more two words are judged to be substitutable in the same linguistic context (i.e. the same location in a sentence), the more synonymous they are in meaning. Church et al. (1994) employ a "lexical substitutability" test in a corpus study of the near-synonyms *ask for*, *request* and *demand*, which produced the same finding: the substitutability of lexical items in the same linguistic context constitutes a good indicator of their semantic similarity. Gries (2001) quantifies the similarity between English adjectives ending in *-ic* or *-ical* (like *economic* and *economical*) on the basis of the overlap between their collocations. 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 and Otani (2010) studies 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)

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 & 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, 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 Learn and Acquire

Concordance enables researchers to compare frequencies of synonymous words. As shown in Table 1, the frequency of *learn* is nearly 3 times of *acquire*.

Table 1. Frequency of *learn* and *acquire* in BNC (per million)

	learn	acquire
Total	18,871	6,712
Per million	168.29	59.83

4.2 The Collocates of Learn and Acquire.

Table 2 and Table 3 list the top 50 left and right collocates of *learn* automatically generated by the software. Table 4 and table 5 list the top 50 left and right collocates of acquire automatically generated by the software.

Table 2. The top 50 left collocates of *learn* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	lessons	217	8.401	26	lot	120	6.413
2	hon.	281	8.322	27	right	241	6.41
3	lesson	130	7.66	28	they	880	6.342
4	children	331	7.475	29	never	167	6.306
5	soon	160	7.145	30	I	2014	6.227
6	child	173	7.097	31	people	294	6.183
7	've	404	6.979	32	student	56	6.164
8	opportunity	107	6.927	33	my	317	6.148
9	teaching	100	6.907	34	need	156	6.114
10	pupils	94	6.866	35	surprise	50	6.112
11	skills	97	6.844	36	will	557	6.11
12	we	963	6.796	37	you	1244	6.102
13	language	120	6.768	38	how	195	6.091
14	surprised	79	6.76	39	she	537	6.026
15	must	279	6.697	40	'd	160	6.024
16	We	328	6.69	41	They	201	6.006
17	have	1475	6.678	42	to	5068	6.005
18	learning	84	6.659	43	has	531	5.993
19	You	339	6.65	44	experience	75	5.969
20	can	733	6.625	45	much	201	5.965
21	students	95	6.605	46	thing	98	5.944
22	quickly	86	6.532	47	having	92	5.934
23	had	1162	6.444	48	should	234	5.932
24	what	511	6.433	49	young	87	5.915
25	learn	70	6.42	50	things	107	5.906

As shown in Table 2, the dominant left collocates of *learn* can be grouped into four categories:

- Abstract nouns: lesson(s), opportunity, teaching, skills, language, experience, thing(s), learning, surprise
- Individual/collective nouns: children, child, pupils, student(s), people
- Personal pronouns: we, you, they, I, my, she
- Auxiliary and modal verbs: ‘ve, have, had, will, ‘d, has, having, must, can, need, should

In addition to the above categories, pronoun, adverb and adjective collocates are also quite salient. Of the 50 collocates there are four adverbs: *soon*, *quickly*, *never* and *to*; 4 interrogative and indefinite pronouns: *what*, *how*, *lot* and *much*; two adjective: *young* and *surprised*. Besides, collocates such as *hon.* and *right* appear in the phrase *(as) my right hon*.

Table 3. The top 50 right collocates of *learn* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	Friend	8.406	8.612	26	play	4.897	6.42
2	how	6.384	8.522	27	speak	5.598	6.398
3	lesson	9.061	8.252	28	disabilities	8.809	6.361
4	skills	7.248	8.058	29	school	4.563	6.311
5	language	6.631	8.026	30	trade	4.923	6.265
6	about	5.519	7.844	31	new	4.005	6.221
7	experience	6.112	7.629	32	swim	7.674	6.22
8	read	6.111	7.466	33	more	3.884	6.218
9	lot	5.629	7.333	34	much	4.03	6.209
10	lessons	8.119	7.298	35	anything	4.522	6.207
11	live	5.878	7.23	36	accept	5.295	6.183
12	from	4.822	7.22	37	craft	7.196	6.154
13	something	5.186	7.166	38	through	3.982	6.138
14	mistakes	8.426	7.143	39	languages	6.606	6.119
15	cope	6.941	6.863	40	art	5.085	6.109
16	techniques	6.443	6.792	41	recognise	6.305	6.103
17	things	4.804	6.704	42	methods	5.352	6.085
18	English	5.084	6.671	43	experiences	6.313	6.078
19	fly	6.869	6.658	44	learning	5.316	6.074
20	quickly	5.498	6.565	45	that	3.605	6.043
21	drive	5.826	6.536	46	ride	6.393	6.006
22	use	4.456	6.525	47	to	3.539	5.991
23	deal	5.127	6.497	48	hard	4.417	5.968
24	write	5.553	6.474	49	understand	4.683	5.96
25	learn	5.758	6.461	50	processes	5.662	5.932

As shown in Table 3, the dominant right collocates of *learn* can be grouped into two categories:

- Abstract nouns: lesson(s), skills, language(s), experience(s), something, mistakes, techniques, things, disabilities, trade, anything, craft, art, methods, learning, processes
- Notional verbs: read, live, cope, fly, drive, use, deal, write, learn, play, speak, swim, accept, recognise, ride, understand

Other collocates such as *about*, *from*, *through*, *how* and *that* have much to do with the grammatical relation which will be analyzed in the next section. Besides, pronoun, adverb and adjective collocates are also salient. Of 50 collocates there are 3 indefinite pronouns: *lot*, *more* and *much*; 2 adverbs: *quickly* and *to*; 2 adjectives: *new* and *hard*.

Table 4. The top 50 left collocates of *acquire* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	newly	95	8.387	26	subsequently	15	5.62
2	recently	95	7.425	27	infection	14	5.617
3	skills	81	7.419	28	gradually	14	5.541
4	knowledge	76	6.915	29	student	19	5.535
5	Newco	24	6.777	30	intent	12	5.531
6	purchaser	25	6.629	31	buyer	13	5.527
7	assets	30	6.493	32	enable	16	5.518
8	company	96	6.251	33	Has	365	5.518
9	skill	23	6.221	34	information	57	5.49
10	definitive	16	6.17	35	ability	21	5.454
11	shares	32	6.165	36	Museum	14	5.427
12	Inc	28	6.149	37	asset	11	5.409
13	TO	24	6.079	38	agreement	24	5.401
14	opportunity	33	6.022	39	process	37	5.401
15	thus	32	5.901	40	means	40	5.326
16	AIDS	15	5.846	41	Soon	30	5.304
17	bidder	12	5.788	42	subsidiary	10	5.281
18	companies	39	5.786	43	collection	16	5.277
19	thereby	15	5.722	44	Corp	14	5.274
20	land	40	5.717	45	had	502	5.273
21	managed	22	5.688	46	volatiles	8	5.27
22	property	28	5.645	47	compulsorily	8	5.265
23	Group	22	5.645	48	person	36	5.245
24	rapidly	17	5.64	49	able	41	5.21
25	students	30	5.633	50	prevent	15	5.202

As shown in Table 4, the dominant left collocates of *acquire* can be grouped into three categories:

- Adverbs: newly, recently, To, thus, thereby, rapidly, subsequently, gradually, soon, compulsorily
- Abstract nouns: skill(s), knowledge, asset(s), shares, opportunity, property, infection, intent, information, ability, agreement, process, means
- Individual/collective nouns: purchaser, company, bidder, companies, student(s), buyer, subsidiary, collection, person

In addition to the above categories, proper noun, notional verb, auxiliary verb and material noun collocates are also quite salient. Of 50 collocates there are 6 proper nouns: *Newco*, *Inc*, *AIDS*, *Group*, *Museum* and *Corp*; 4 notional verbs: *managed*, *enable*, *able* and *prevent*; 2 auxiliary verbs: *has* and *ha*d; 2 material nouns: *land* and *volatiles*.

Table 5. The top 50 right collocates of *acquire* in BNC

Rank	Collocates	Freq	logDice	Rank	Collocates	Freq	logDice
1	skills	163	8.428	26	Museum	21	6.012
2	knowledge	159	7.98	27	language	46	5.978
3	reputation	76	7.912	28	properties	20	5.953
4	shares	103	7.851	29	information	78	5.942
5	assets	62	7.54	30	citizenship	14	5.932
6	title	58	7.048	31	ownership	18	5.924
7	status	55	6.87	32	competence	15	5.916
8	qualifications	31	6.851	33	sufficient	23	5.909
9	land	83	6.771	34	software	25	5.875
10	stake	28	6.71	35	deficiency	13	5.857
11	skill	32	6.698	36	meaning	25	5.821
12	expertise	26	6.521	37	immune	13	5.798
13	understanding	40	6.376	38	through	143	5.767
14	Target	18	6.363	39	interest	55	5.763
15	taste	26	6.338	40	asset	14	5.756
16	infection	23	6.334	41	dispose	12	5.753
17	Newco	17	6.28	42	necessary	40	5.738
18	additional	32	6.277	43	company	67	5.732
19	property	43	6.264	44	goods	26	5.708
20	rights	42	6.239	45	habit	14	5.692
21	Inc	29	6.2	46	momentum	12	5.685
22	syndrome	16	6.085	47	power	56	5.665
23	premises	22	6.083	48	wealth	16	5.665
24	significance	23	6.061	49	weapons	16	5.64
25	during	86	6.038	50	collection	20	5.599

As shown in Table 5, the dominant left collocates of *acquire* can be grouped into two categories:

- Abstract nouns: skill(s), knowledge, reputation, asset(s), title, status, qualifications, expertise, understanding, taste, infection, rights, syndrome, significance, language, information, citizenship, ownership, competence, deficiency, meaning, interest, habit, momentum, power, wealth
- Individual/collective nouns: shares, stake, property, premises, properties, software, company, goods, weapons, collection

In addition to the above categories, adjective, proper noun and preposition collocates are also quite salient. Of the 50 collocates there are 4 adjectives: *additional*, *sufficient*, *immune* and *necessary*; 4 proper nouns: *Target*, *Newco*, *Inc* and *Museum*; 2 prepositions: *during* and *through*.

4.3 The Syntactic Patterns of Learn and Acquire

The syntactic patterns of the two verbs are based on the Word Sketch function of SkE. In order to present a fine-grained comparison, I summarized the 18 patterns of *learn* and 14 patterns of *acquire* in Table 6 and Table 7. In the first example of Table 6, the underlined word *beginners* functions as the subject of *learn*.

Table 6. The syntactic behavior of *learn* in BNC

Categories	Freq	Score	Example
subject	2282	2.9	***Beginners*** can also ***learn*** in other resorts.
object	4989	3.4	She's ***learned*** a bitter ***lesson*** yesterday,
modifier	2664	0.6	Alison ***soon learned*** my style.
pp_from-p	962	10.9	Students ***learn*** best ***from*** their own mistakes;
pp_about-p	771	33.9	so the course is ***learning about*** the nature of mechanical
and/or	374	0.2	helped Diana to listen ***and learn*** from counselling sessions
pp_in-p	372	1.0	'we shall all know what they are all **learning in** school',
pp_of-p	229	0.3	We first ***learnt of*** its existence in May,
pp_by-p	157	1.4	Facts to be ***learned by*** rote are often best assimilated just
pp_at-p	137	1.5	although it can be ***learned at*** a very early age in the nest
pp_through-p	95	6.1	enables members to ***learn through*** shared experience.
np_adj_comp	74	3.3	Well, you ***learn*** something ***new*** every day.
part_intrans	71	0.5	your child in his/her efforts to ***learn about*** and cope with life.
pp_to-p	60	0.3	As the Royals have ***learned to*** their cost the law is so lax that
pp_on-p	60	0.4	ensure that they ***learn on*** the job and produce an effective
pp_for-p	49	0.3	This was when I ***learnt for*** the first time how experts conduct
part_trans	26	0.3	You had to ***learn*** them ***off*** by heart.
pp_over-p	23	1.5	but one thing I've ***learned over*** the years is that

Table 7. The syntactic pattern of *acquire* in BNC

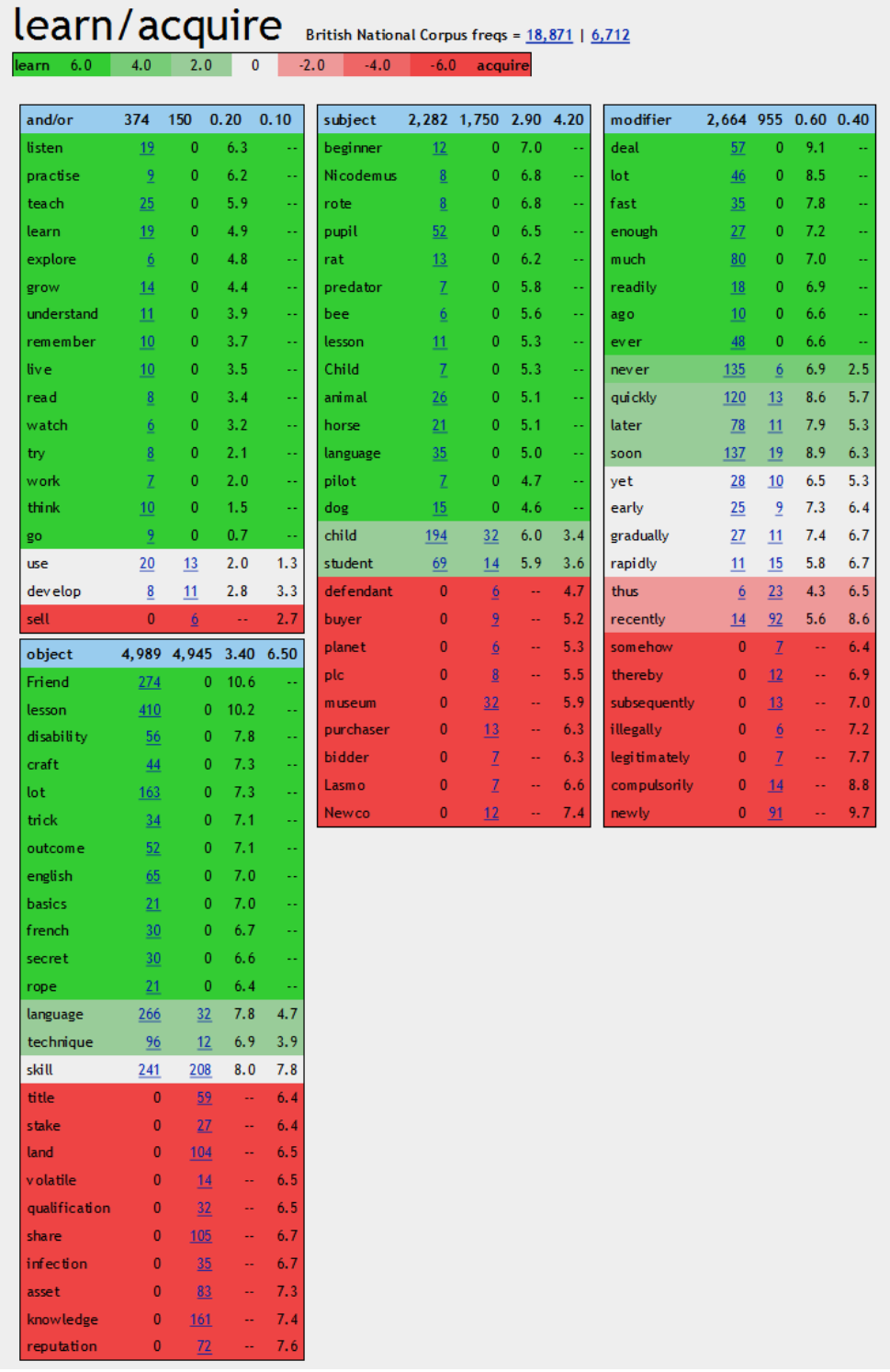
Categories	Freq	Score	Example
subject	1750	4.2	The British ***Museum acquired*** some of these pieces knowingly,
object	4945	6.5	the need for the traditional archivist to ***acquire*** new ***skills***
modifier	955	0.4	of an international market for their ***newly acquired*** product
pp_by-p	372	6.4	Four of those ***acquired by*** the National Museums are of Scottish
pp_in-p	163	0.9	comprised in the skills ***acquired in*** the course of employment
and/or	150	0.1	Interviewing skills can be ***acquired and*** developed by placing
pp_through-p	70	8.5	However cases of AIDS ***acquired through*** heterosexual contact
np_adj_comp	68	5.8	so that they can ***acquire*** the skills ***necessary*** to collect and use
pp_during-p	57	13.2	if the information ***acquired during*** pre-exposure is to be fully
pp_for-p	56	0.6	which might be ***acquired for*** investment purposes
pp_at-p	36	0.7	an ideal strategic weapon ***acquired at*** a modest cost
pp_as-p	28	1.4	Common law rights are ***acquired as*** a result of custom and
pp_on-p	26	0.4	use the IT skills ***acquired on*** their Advanced Courses in the
pp_over-p	22	2.7	have proven very difficult to ***acquire over*** the two-year period

It has to be noted that although the syntactic patterns of the two verbs are similar in many ways, there also exist apparent differences, which can be easily shown when using Sketch-Diff function of SkE.

4.4 Direct Comparison of Lexical and Grammatical Collocates

The Sketch-Diff function of SkE allows users to visually compare and contrast synonymous words according to

their salient collocational context. Figure 1 is part of the result when clicking 'Show Diff'. In the figure, the greener a word is, the more closely it relates to *learn*. The redder a word is, the more closely it relates to *acquire*. For example, it is more usual to say *pupil learns* than *buyer acquires*.



learn/acquire British National Corpus freqs = 18,871 | 6,712

learn 6.0 4.0 2.0 0 -2.0 -4.0 -6.0 acquire

and/or	374	150	0.20	0.10
listen	19	0	6.3	--
practise	9	0	6.2	--
teach	25	0	5.9	--
learn	19	0	4.9	--
explore	6	0	4.8	--
grow	14	0	4.4	--
understand	11	0	3.9	--
remember	10	0	3.7	--
live	10	0	3.5	--
read	8	0	3.4	--
watch	6	0	3.2	--
try	8	0	2.1	--
work	7	0	2.0	--
think	10	0	1.5	--
go	9	0	0.7	--
use	20	13	2.0	1.3
develop	8	11	2.8	3.3
sell	0	6	--	2.7

object	4,989	4,945	3.40	6.50
Friend	274	0	10.6	--
lesson	410	0	10.2	--
disability	56	0	7.8	--
craft	44	0	7.3	--
lot	163	0	7.3	--
trick	34	0	7.1	--
outcome	52	0	7.1	--
english	65	0	7.0	--
basics	21	0	7.0	--
french	30	0	6.7	--
secret	30	0	6.6	--
rope	21	0	6.4	--
language	266	32	7.8	4.7
technique	96	12	6.9	3.9
skill	241	208	8.0	7.8
title	0	59	--	6.4
stake	0	27	--	6.4
land	0	104	--	6.5
volatile	0	14	--	6.5
qualification	0	32	--	6.5
share	0	105	--	6.7
infection	0	35	--	6.7
asset	0	83	--	7.3
knowledge	0	161	--	7.4
reputation	0	72	--	7.6

subject	2,282	1,750	2.90	4.20
beginner	12	0	7.0	--
Nicodemus	8	0	6.8	--
rote	8	0	6.8	--
pupil	52	0	6.5	--
rat	13	0	6.2	--
predator	7	0	5.8	--
bee	6	0	5.6	--
lesson	11	0	5.3	--
Child	7	0	5.3	--
animal	26	0	5.1	--
horse	21	0	5.1	--
language	35	0	5.0	--
pilot	7	0	4.7	--
dog	15	0	4.6	--
child	194	32	6.0	3.4
student	69	14	5.9	3.6
defendant	0	6	--	4.7
buyer	0	9	--	5.2
planet	0	6	--	5.3
plc	0	8	--	5.5
museum	0	32	--	5.9
purchaser	0	13	--	6.3
bidder	0	7	--	6.3
Lasmo	0	7	--	6.6
Newco	0	12	--	7.4

modifier	2,664	955	0.60	0.40
deal	57	0	9.1	--
lot	46	0	8.5	--
fast	35	0	7.8	--
enough	27	0	7.2	--
much	80	0	7.0	--
readily	18	0	6.9	--
ago	10	0	6.6	--
ever	48	0	6.6	--
never	135	6	6.9	2.5
quickly	120	13	8.6	5.7
later	78	11	7.9	5.3
soon	137	19	8.9	6.3
yet	28	10	6.5	5.3
early	25	9	7.3	6.4
gradually	27	11	7.4	6.7
rapidly	11	15	5.8	6.7
thus	6	23	4.3	6.5
recently	14	92	5.6	8.6
somehow	0	7	--	6.4
thereby	0	12	--	6.9
subsequently	0	13	--	7.0
illegally	0	6	--	7.2
legitimately	0	7	--	7.7
compulsorily	0	14	--	8.8
newly	0	91	--	9.7

Figure 1. Comparison of *learn* and *acquire* in terms of collocational patterns

Apparently, despite that the two verbs *learn* and *acquire* share a number of syntactical patterns, the collocates in each pattern differ considerably. In the 'and/or' pattern, the collocation tokens for *learn* are 374 and 150 for

acquire, indicating that there are more words used in 'and/or' pattern with *learn*. Many words frequently collocates with *learn* but never used with *acquire* (such as *listen*, *practice*, *teach*, *learn*, *explore*, *grow*, *understand*, *remember*, *live*, *read*, *watch*, *try*, *work*, *think*, and *go*). On the other hand, *sell* occurs 6 times with *acquire*, but there is no occurrence of *sell* with *learn*.

In the 'modifier' pattern, the collocation tokens for *learn* are 2664 and 955 for *acquire,* indicating that there are more words used as modifier of *learn*. Some words only collocate with *learn* (such as *deal*, *lot*, *fast*, *enough*, *much*, *readily*, *ago* and *ever*), some only collocate with *acquire* (such as *somehow*, *thereby*, *subsequently* and *illegally*), some others can be used both with *learn* and *acquire* (such as *never*, *quickly*, *later*, *soon*, *yet*, *early*, *gradually*, *rapidly*, *thus* and *recently*). As is shown in Fig 1, some words are more likely to collocate with *learn* than with *acquire* (such as *never*, *quickly*, *later*, *soon*, *yet*, *early* and *gradually*), while some others are more likely to collocate with *acquire* than with *learn* (*rapidly*, *thus* and *recently*).

In the 'object' pattern, the collocation tokens for *learn* are 4989 and 4945 for *acquire,* indicating that there are similar numbers of words used as objects of *learn* and *acquire*. Some words only collocate with *learn* (such as *Friend*, *lesson*, *disability*, *craft*, *lot*, *trick*, *outcome*, *English*, *basics*, *French*, *secret and rope*), some only collocate with *acquire* (such as *title*, *stake*, *land*, *volatile*, *qualification*, *share*, *infection*, *asset*, *knowledge and reputation*), some others can be used both with *learn* and *acquire* (such as *language*, *technique and skill*). As is shown in Fig 1, some words are more likely to collocate with *learn* than with *acquire* (such as *never*, *quickly*, *later*, *soon*, *yet*, *early* and *gradually*), while some others are more likely to collocate with *acquire* than with *learn* (*rapidly*, *thus* and *recently*). Something worth noting is that the subjects of *learn* is more about knowledge or skill (such as *lesson*, *craft*, *trick*, *English*, *basics*, *French*, *rope*, *language*, *technique* and *skill*), while the subjects of acquire are more about qualification or asset (such as *title*, *qualification*, *reputation*, *stake*, *land*, *share* and *asset*).

5. Conclusion

In view of its importance and intricacy, researching synonymy is a crucial task in the field of lexical semantics. This paper has 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 *learn* and *acquire*.

This study has a number of pedagogical implications. First, studies in second language acquisition show that native-speakers memorize not only words in isolation, but also chunks of words. These chunks are viewed as the building blocks of language. They are available to speakers as ready-made or prefabricated units, and therefore contribute to conferring fluency and naturalness of their utterances. Thus, if EFL teachers aim to help their students to achieve a great amount of fluency and accuracy, they may hope students to use the collocational and colligational patterns from Table 1 to 7. Second, there exist a huge amount of synonyms in English, therefore it would be unlikely for teachers to teach each pair of them to students. It might be more promising to teach students how to use SkE to conduct their own research.

Acknowledgement

The research was supported by Innovative School Project in Higher Education of Guangdong, China (GWTP-BS-2014-13) and Foundation for Outstanding Young Teachers, Guangdong Province (GWTP-SY-2014-08).

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. http://dx.doi.org/10.1093/ijl/8.2.85

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. http://dx.doi.org/10.1017/CBO9780511804489

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.

Church, K. W., Gale, W., Hanks, P., Hindle, D., & Moon, R. (1994). Lexical substitutability. In B. T. S. Atkins, & A. Zampolli (Eds.), *Computational Approaches to the Lexicon* (pp. 153-157). Oxford: Oxford University Press.

DiMarco, C., Hirst, G., & Stede, M. (1993). The semantic and stylistic differentiation of synonyms and near-synonyms. In *AAAI Spring Symposium on Building Lexicons for Machine Translation* (pp. 114-121). Stanford, CA, March.

Divjak, D., & Gries, S. Th. (2006). Ways of trying in Russian: Clustering behavioral profiles. *Journal of Corpus Linguistics and Linguistic Theory, 2*(1), 23-60. http://dx.doi.org/10.1515/cllt.2006.002

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* (pp. 168-205). London: Longmans.

Gilquin, G. (2003). Causative 'get' and 'have': So close, so different. *Journal of English Linguistics, 31*(2), 125-148. http://dx.doi.org/10.1177/0075424203031002002

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

Gries, S. Th. (2001). A corpus-linguistic analysis of -ic and -ical adjectives. *ICAME Journal, 25*, 65-108.

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

Hanks, P. (1996). Contextual dependency and lexical sets. *International Journal of Corpus Linguistics, 1*(1), 75-98. http://dx.doi.org/10.1075/ijcl.1.1.06han

Hoey, M. (2005). *Lexical Priming: a new theory of words and language*. Routledge. http://dx.doi.org/10.4324/9780203327630

Hu, C., & Yang, B. (2015). Using Sketch Engine to investigate synonymous verbs. *International Journal of English Linguistics, 5*(4), 29-41. http://dx.doi.org/10.5539/ijel.v5n4p29

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524773

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

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). "The Sketch Engine: Ten years on". *Lexicography, 1*(1), 7-36. http://dx.doi.org/10.1007/s40607-014-0009-9

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. http://dx.doi.org/10.1075/ijcl.15.1.03liu

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. http://dx.doi.org/10.1075/ijcl.17.2.03liu

Longman Synonym Dictionary. (1986). L. Urdang (Ed.), Harlow, England: Longman.

Merriam Webster's Dictionary of Synonyms. (1984). Merriam-Webster (Ed.). Massachusettes, U.S.A.: Merriam Webster.

Miller, G. A., & Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes, 6*(1), 1-28. http://dx.doi.org/10.1080/01690969108406936

Palmer, H. E. (1933). *Second interim report on English collocations*. Tokyo: Kaitakusha.

Partington, A. (1998). *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam, Netherlands: John Benjamins. http://dx.doi.org/10.1075/scl.2

Peirsman, Y., Geeraerts, D., & Speelman, D. (2015). The corpus-based identification of cross-lectal synonyms in pluricentric languages. *International Journal of Corpus Linguistics, 20*(1), 55-81. http://dx.doi.org/10.1075/ijcl.20.1.03pei

Sinclair, J., Jones, S., Daley, R., & Krishnamurthy, R. (2004). *English Collocational Studies: The OSTI Report*. London: Continuum.

Xiao, R., & T, McEnery. (2006). Collocation, semantic prosody, and near synonymy: A crosslinguistic perspective. *Applied Linguistics, 27*(1), 103-129. http://dx.doi.org/10.1093/applin/ami045