

# Corpus-aided Business English Collocation Pedagogy: An Empirical Study in Chinese EFL Learners

Lidan Chen<sup>1</sup>

<sup>1</sup> School of English for International Business, Guangdong University of Foreign Studies, Guangzhou, China

Correspondence: Lidan Chen, School of English for International Business, Guangdong University of Foreign Studies, 2 Baiyun Ave (N), Baiyun District, GZ 510420, China. E-mail: cld4writing@126.com

Received: July 28, 2017 Accepted: August 16, 2017 Online Published: August 18, 2017

doi: 10.5539/elt.v10n9p181 URL: <http://doi.org/10.5539/elt.v10n9p181>

## Abstract

This study reports an empirical study of an explicit instruction of corpus-aided Business English collocations and verifies its effectiveness in improving learners' collocation awareness and learner autonomy, as a result of which is significant improvement of learners' collocation competence. An eight-week instruction in keywords' collocations, with the help of AntConc and self-constructed Business English Pedagogical Corpus combined with COCA general corpus and Wikipedia corpus, was imparted to 23 undergraduate learners majoring in Business English in Guangdong University of Foreign Studies. They took the collocation competence pre-test and post-test before and after the teaching experiment which was phased into two themes and submitted learning reflective journals at the end of each theme instruction and answered a questionnaire at the final end. The data from the tests, reflective journals and questionnaire collaboratively suggest that given appropriate guidance EFL Business English learners can take a more active role in raising their collocation awareness and developing learner autonomy and thus improve their collocation competence significantly. The results from the test analysis also indicate that the corpus-aided Business English collocation pedagogy is proved to be more effective for intermediate and advanced level learners rather than lower level ones. The findings have pedagogical implications for EFL Business English instructors and learners.

**Keywords:** corpus-aided pedagogy, Business English, learner autonomy, collocation awareness, collocation competence

## 1. Introduction

An explosion of studies devoted to various aspects of the use of corpora in language learning has been addressed with the development of corpus linguistics. Corpus pedagogy has been more and more often used in English as Foreign Language (EFL) writing and vocabulary instruction. A central idea developed from corpus pedagogy is the collocational syllabus since the single most important task facing language learners is acquiring sufficiently large prefabricated chunks and self-evidently; teaching collocation should be a top priority in every language course (Lewis, 2000). As Lewis (2012) put it, collocations describe the way individual words co-occur with others. In most cases learners will already know many individual words, but when they produce in a written or oral form, they may need more about "putting them together". In other words, if words occur together, learners need to notice that co-occurrence, whose pattern is an important perspective of deep knowledge of words. Therefore, it becomes natural for collocation to assume an important syllabus-generating role (Lewis, 2012). However, it is found (Sun, 2008) that in Chinese setting, because of linguistic and cultural difference, English learners commit collocation errors largely which reduces the naturalness and idiomaticity of English. Based on this, the present study aims to apply corpus resources and corpus techniques to Business English pedagogy in Chinese setting to enhance the collocation competence of EFL learners, which both tentatively enriches corpus-based pedagogy and Business English teaching methods.

## 2. Rationale for Explicit Corpus-aided Pedagogy in Business English

The twofold rationale for corpus-aided pedagogy in Business English is authenticity of teaching materials and abundance of language resources. One of the challenges of EFL learners in China, which is member in the Expanding Circle of World Englishes, is the difficulty to obtain a genuine English context. In China, studies are mostly conducted in classroom under a formal context, not in a natural English-speaking context. Corpus-aided pedagogy allows learners to access large amounts of authentic materials and multiple examples of a particular

feature in the form of concordance lines (Gilquin & Granger, 2010). What's more, Allan (2009) stated the specific potential of corpus-aided lexical pedagogy for lexical learning, "it is particularly helpful in that it gives learners multiple exposures to words in context, offering potential for deepening word knowledge through the information provided about collocations, contextual behaviour, and register. It would appear to be a valuable explicit "focus on form" technique." This key-word-in-context pattern, as such, can highlight word co-occurrence in a natural language context. This process will induce learners to arrive at descriptions of certain language features, collocations, which are more accurate and practical than those stated in non-corpus-informed teaching materials (Johns, 1991).

### 3. Literature Review

Many attempts on corpus pedagogy in English teaching with a focus on collocation have illustrated its rationale and suggested models to follow in a L1 learning environment since the proposal of the term "collocation" in 1957 by Sinclair. According to Poole (2016), Data-Driven Learning for direct in-class corpus study, the corpus-based CoBuild dictionaries and materials, and the concordancing method are early developments in corpus pedagogy, after which there has been explosion of studies in this regard with the teaching and learning of vocabulary through corpus consultation most often the focus. The publication of *Teaching collocation: Further developments in the lexical approach* by Lewis in 2000 was a big step forward in that illustrates theoretical background of lexical approach, and a wealth of classroom practices inclusive of design and implementation in teaching collocations in L1 context, from which the present study draws on a lot ideas. However, challenged for its bottom-up practice with little attention to context, corpus pedagogy has increasingly shifted to more discourse and genre-specific approaches, resulting in deeper insights into particular genres of EAP and ESP interest and more comprehensive explanations of the pedagogical implications (Poole, 2016).

Some overseas researches on corpus-aided pedagogy in Business English have been addressed, most of which are about semantic prosody in Business English lexis in L1 context. Nelson (2006) examined the semantic associations of business words to show how they interact in a business environment and discussed the pedagogical consequences. Walker (2011) used two case studies to show how corpus linguistics can be used to help in the teaching of Business English lexis from the perspective of semantic prosody based on the questions about slight but significant differences in the meaning of the collocates of the node word raised by adult Business English learners.

As for corpus-aided Business English pedagogy in Chinese setting, there has been some literature with regard to writing (Wang & Huang, 2011; Zhang, 2012; Chen, 2015), lexis (Wang, 2009; Hu, 2011 & 2012; Shi, 2012), speaking (Du, 2013; Shen, 2016) and translation (Wei, Zhang, & Zhong, 2014) instructions. Hu (2011, 2012) detailed the rationale and implementation of corpus consultation in Business English lexis teaching in China tentatively. But among all these relevant studies, few empirical studies have been conducted with a teaching experiment in the classroom to prove the effectiveness of corpus-aided collocation pedagogy in Business English classroom.

### 4. Research Purpose and Research Questions

Informed by corpus findings in L1 pedagogical context and inspired by L2 corpus consultation in Business English studies in Chinese setting, the present article briefs the teaching details and reports findings of an empirical study that features corpus-aided collocation pedagogy in L2 undergraduate Comprehensive Business English Course. The approach illustrated here provides a model for incorporating corpus consultation into Business English classroom as it details how KWIC and collocates based on AntConc and self-constructed Business English Pedagogical Corpus (BEPC), COCA general corpus and Wikipedia corpus can be effectively introduced into L2 undergraduate Business English classroom to develop a model for teaching collocation in Business English classroom to strengthen both learners' collocation competence. And two important mediating roles in the process are hypothesized to be learners' collocation awareness and learner autonomy.

The following research questions served as a guide in the present study:

- 1) Does corpus-aided approach raise EFL Business English undergraduates' collocation awareness significantly?
- 2) Does corpus-aided approach bring EFL Business English undergraduates a step closer to antonymous learners?
- 3) Does corpus-aided approach develop learners' collocation competence significantly in a short term? Is there any difference in the effects on learners of different English proficiency?

## 5. Research Methodology and Procedures

### 5.1 Participants

23 undergraduate learners majoring in Business English participated in the present research; they were further divided into three subgroups—advanced, intermediate and lower levels by 30%, 40% and 30%, according to their English proficiency level based on their marks of Comprehensive Business English Course of Year 1 as well as their marks of English subject in College Entrance Examination, both of which each student had individually taken before. Since the latter record was relatively old though comprehensive enough for English proficiency level and the former was more about their Business English proficiency level; consideration of these two can best illustrate their Business English proficiency at the starting point. In light of these two marks, 7 learners were classified as advanced, 9 as intermediate and 7 as lower level.

### 5.2 Instruments

The present study is an empirical study with quantitative and qualitative analysis of collected data. Embedded in second year undergraduate course Comprehensive Business English, an 8-week corpus-aided Business English instruction with focus on Business English collocations was conducted. A pre-test and a post-test, an online questionnaire and learning reflective journals of the instructed learners were tools used to generate and collect quantitative and qualitative data to answer the research questions.

### 5.3 Teaching Resources

The present corpus-aided teaching experiment offers the teacher and learners the opportunity to work with corpus data, which come from teacher-oriented BEPC, COCA general corpus, COCA Wikipedia corpus, and Learners Corpus: learners' essay corpus.

Business English Pedagogical Corpus (621,000 tokens) is composed of two textbooks recommended to all the learners in this course in the university: *Business English: A Comprehensive Course* (Higher Education Press, 2009) and *Business Essentials* (Pearson, 2013). As the learners were required to read or listen to the texts which constitute the corpus, they were supposed to understand the context and co-text of specific features chosen for study. The corpus data were provided by the teacher, accessed by the learner under the teacher's guidance, or accessed by the learner autonomously. In some cases the corpus data were filtered by the teacher before classroom use, typically by selecting the data which were suitable for the learners. When given the corpus, learners could use AntConc to carry out some preliminary investigations like keyword list, collocates and concordance.

The learners were instructed to navigate the Corpus of Contemporary American English (COCA) general corpus (<http://corpus.byu.edu/coca/>) with the focus on collocations of high frequency keywords and conduct self-access learning on any lexical items they were interested in. In other words, learners were guided to look at lexis based on certain genre based on cross-discipline search of COCA general corpus.

Wikipedia corpus (<http://corpus.byu.edu/wiki/>) from COCA online presents an update view of online resources from Wikipedia. Responding to learners' feedback about insufficient data from pedagogical corpus, the teacher introduced it, in the teaching of the second theme and lectured how to compose mini corpus of certain business topic according to specific needs, on which similar investigations to COCA general corpus could be conducted and wordlist and collocation list on assigned business topics could be retrieved for self-access learning. The following procedures illustrate how to administer Wikipedia corpus.

Step 1: Open <http://corpus.byu.edu/wiki/>, register and log in and click on "Create corpus" button

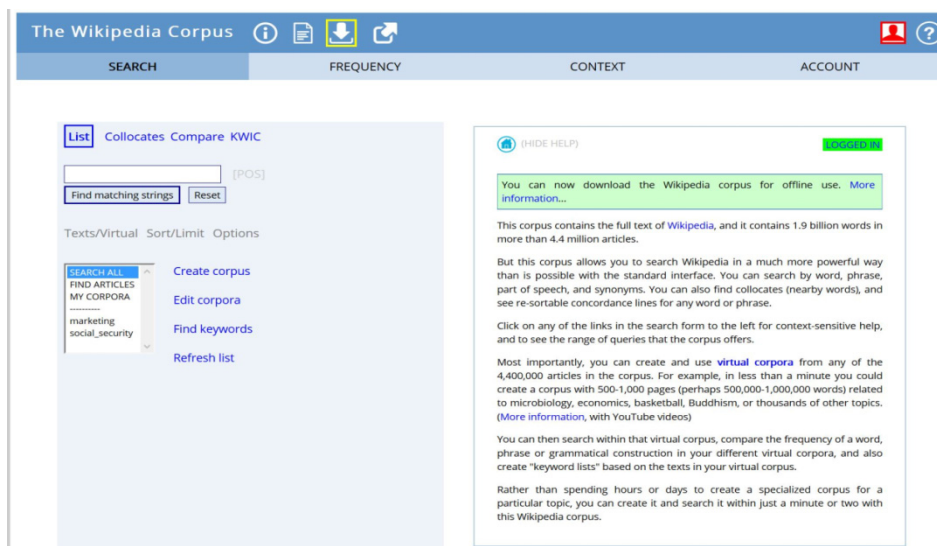


Figure 1. Wikipedia corpus home page

Step 2: Search for the title word “social security” and click on “submit” button to begin search

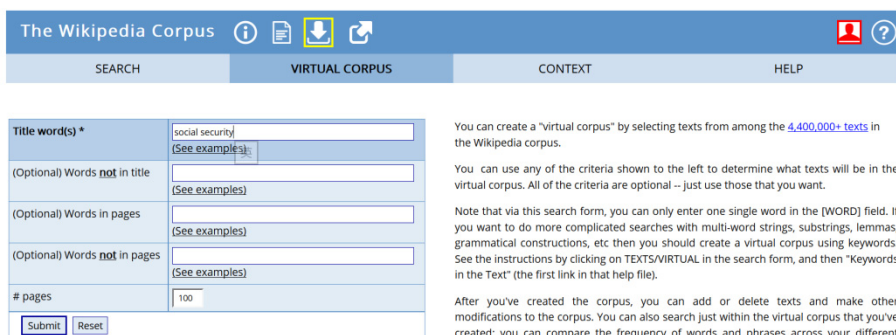


Figure 2. Construct a corpus by searching the title word “social security”

Step 3: Name the self-constructed corpus and click on “submit” button to save it, and then it will be listed in “virtual corpus”

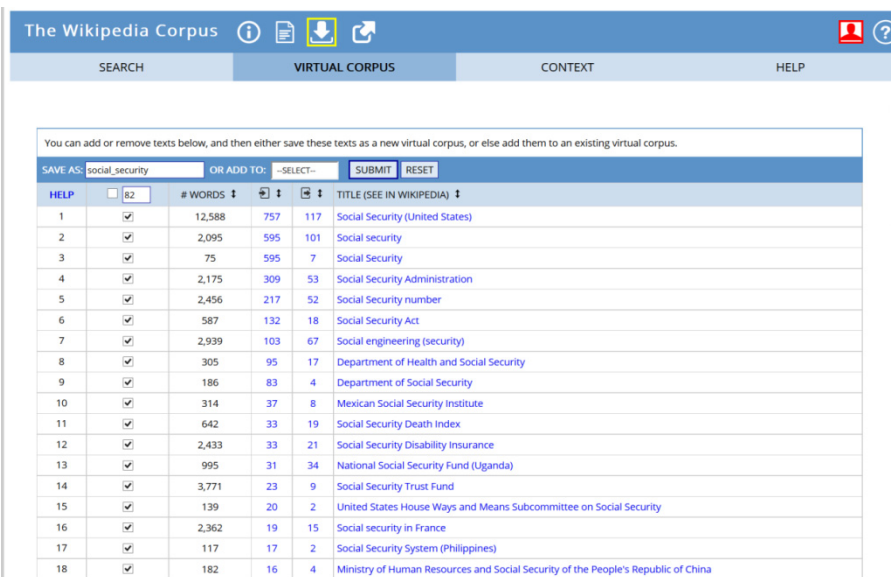


Figure 3. Texts retrieved with “social security” as or in the title

The Wikipedia Corpus interface shows a navigation bar with tabs: SEARCH, VIRTUAL CORPORA, CONTEXT, and OVERVIEW. Below the navigation bar, there is a section titled "MY VIRTUAL CORPORA" with a table listing two corpora.

HELP		LIST NAME	# ARTICLES	# WORDS	FIND KEYWORDS	CREATED
1		SOCIAL_SECURITY	82	72,475	NOUN VERB ADJ ADV N+N ADJ+N	163 d
2		MARKETING	100	187,859	NOUN VERB ADJ ADV N+N ADJ+N	163 d

Figure 4. A list of self-constructed virtual corpora

Step 4: Retrieve by clicking on any wordlist or collocation list from the virtual corpus

The Wikipedia Corpus interface shows the "KEYWORDS" tab. A table displays the "N+N" collocation list for the "SOCIAL\_SECURITY" corpus. The table includes columns for word frequency, number of texts, and expected frequency.

HELP	WORD (CLICK TO SEE)	FREQ	# TEXTS	SPECIFIC	ALL WIKIPEDIA	EXPECTED
1	PAYROLL TAX	74	6	2,915.0	648	0.0
2	TRUST FUND	68	6	617.1	2,813	0.1
3	RETIREMENT AGE	48	3	744.8	1,645	0.1
4	SECURITY SYSTEM	30	18	317.0	2,416	0.1
5	HEALTH INSURANCE	29	9	103.5	7,154	0.3
6	RETIREMENT BENEFIT	24	4	1,373.6	446	0.0
7	TAX RATE	22	6	97.5	5,761	0.2
8	DISABILITY BENEFIT	20	6	1,904.9	268	0.0
9	INSURANCE PROGRAM	17	8	477.4	909	0.0
10	RETIREMENT SYSTEM	15	6	463.5	826	0.0
11	PENSION PLAN	14	5	299.0	1,195	0.0
12	INCOME TAX	13	5	29.0	11,461	0.4
13	TAX INCREASE	12	4	212.7	1,440	0.1
14	UNEMPLOYMENT BENEFICIARY	12	1	23,562.5	13	0.0
15	HEALTH CARE	10	7	7.7	33,364	1.3
16	JOB SEEKER	10	1	545.4	468	0.0
17	INTEREST RATE	9	4	27.0	8,522	0.3
18	GOVERNMENT EMPLOYEE	9	4	114.6	2,005	0.1
19	BENEFIT PAYMENT	9	4	1,867.8	123	0.0

Figure 5. "N+N" collocation list from "social security" virtual corpus

And then, all the self-constructed corpora will be in Wikipedia Corpus interface for further investigation.

The Wikipedia Corpus interface shows the "FREQUENCY" tab. The "Collocates Compare Concordances" section is active, displaying a search box with the word "pension" and a list of options: "List", "Collocates", "Compare", and "Concordances". Below the search box, there are buttons for "Find matching strings" and "Reset". To the right, there is a "New interface" button and a "Create corpus" button. The interface also includes a "Texts/Virtual Sort/Limit Options" section and a "SEARCH ALL FIND ARTICLES MY CORPORA" section.

Figure 6. Wikipedia corpus with self-constructed virtual corpora

Learners Corpus is also an important learning resource in the present study, which was constructed by highly-motivated learners without prompt. They collected all the essays written in the Business English Writing course and explored it by use of AntConc to get interesting findings about their essays. By investigating it across time, comparing to sample composition corpus or peer composition corpus, certain linguistic patterns and features in their own written output would be discovered.

## 5.4 Procedures

### 5.4.1 Pre-test

The learners were required to sit for a pre-test on collocation competence of 10 business words from pedagogical corpus composed of two textbooks used in this course. The words examined were chosen based on the keyness and frequency based on the typology of business lexis of Nelson (2000) (Note 1). The collocation pre-test was designed to measure learners' mastery of receptive and productive collocation knowledge of key words by use of multiple choices, gap filling with the initial letter clue and discourse completion test. The four increasingly difficult tasks are designed to test learners' collocation knowledge of the key words from four perspectives: recognition of collocations in certain business context (receptive knowledge), distinguish associated/similar collocations (receptive knowledge), produce collocation with a language clue (productive knowledge), and produce collocations without a language clue in certain business context (productive knowledge).

### 5.4.2 Learners Training

After the pretest, 2-week session on learners training about corpus-aided pedagogy was conducted in order to get them into picture. It equipped the learners with basics of corpus-aided approach so that they were highly motivated to be exposed to corpus-retrieved teaching materials and adequately able to harness the corpus skills. It has been instructed that collocation is integrated as an organizing principle within syllabus (Lewis, 2012) and the Present-Practice-Produce paradigm is rejected, in favor of a paradigm based on the Observe-Hypothesis-Experiment cycle (Lewis, 2012). They were also inducted into the skills of software AntConc and the online platform COCA general corpus with focus on basic research on collocates, concordance, and keyword list, as well as basics of corpora compilation. Both of these were listed as essential tools for corpus-aided directed learning in class and autonomous learning after class.

### 5.4.3 Weekly Sessions

During the 6 weekly sessions on two business themes: marketing and social security, learners were requested to take part in classroom activities and submitted learning reflective journals through an online corpus-based platform (<http://www.pigai.org/>). Weekly sessions were phased into 2 business-theme schemes in the research. Teacher's feedback, business collocation practice and learners' reports of their discoveries in the assignments composed the general model of each weekly session as in Table 1. Immediately after weekly session, the learners were required to submit their learning reflections and received the teacher's timely feedback.

Table 1. Model of weekly sessions

Activities	Tasks
In-class	<ul style="list-style-type: none"> <li>-Feedback on learners' reflective journals of last week</li> <li>-Corpus-aided activities</li> <li>-Theoretical framework of business topic based on retrieved high frequency keywords and their collocates</li> <li>-Collocation exercises based on concordance lines and collocates of keywords (gap-filling and Q&amp;A, translation and writing, etc.)</li> <li>-Learners' report on autonomous learning</li> </ul>
After-class	<ul style="list-style-type: none"> <li>-Learners' submission of learning reflective journals and accomplishment of weekly autonomous learning tasks</li> <li>-Teacher's responding to learners' reflective journals</li> </ul>

The sample corpus consultation task below illustrates how the teacher included questions and prompts along with the corpus data to guide learners' discoveries. In this case, the learners used the corpus data to investigate collocations of "cut" with focus on business contexts.

Task 1 Read the following incomplete sentences and answer the questions afterwards.

- 1) Chinese government \_\_\_\_\_ the current tariff on skin-care products to spur consumption.
- 2) In Hong Kong, import licensing requirements have been \_\_\_\_\_.
- 3) ... \_\_\_\_\_ state trading monopolies in order to vitalize the market.

Questions:

1. Could you fill in the gap with a verb?
2. Can you translate them into Chinese?
3. How can you find out more business words after the missing words by use of any corpus tool introduced in this course?

Following this weekly session model what we had attempted to do in the teaching of the first theme is demonstrated in Table 2.

Table 2. Teaching plan of the first theme

Duration	Objectives	Tasks
2 weeks	-Give a general picture of theoretical basics of marketing based on keywords from BEPC	-Presenting the mind map of marketing as a discipline based on retrieved keywords on marketing and their collocates -Listing out similar/associated business phrases retrieved from pedagogical corpus for comparison
1 week	-Induce learners to master important collocates of high frequency business keywords about marketing -Motivate learners to conduct autonomous collocation learning by use of COCA general corpus and BEPC	-Gap-filling and Q&A based on collocation exercises based on concordance lines from BEPC and COCA general corpus -Lecturing basics of business diagrams and tasks on self-access learning of collocation structures in graph description based on corpus resources -Learners' report of assigned tasks on key words and phrases in textbook as well as graph description

Responding to learners' reflective journals, the teacher made changes in the classroom activities accordingly in the teaching of the second theme presented in Table 3 to make the best out of corpus-aided approach, with simplified arrangement of lexis activities and more highly autonomous learning on COCA Wikipedia corpus.

Table 3. Teaching plan of the second theme

Duration	Objectives	Tasks
2 weeks	-Instruct how to administer Wikipedia corpus -Motivate learners to discuss about hot topics in social security based on corpus consultation	-Learners' construction of social security mini corpus in Wikipedia corpus -Learners' retrieving words and collocate list from learners' self-constructed mini corpus -Learners' group discussion and report on 3 hot topics about social security based on retrieved keywords and collocations
1 week	-Induce learners to master important collocates of high frequency business keywords about social security -Motivate learners to conduct autonomous collocation learning by use of COCA, general corpus, Wikipedia corpus and BEPC with AntConc	-Gap-filling and Q&A based on collocation exercises based on concordance lines from BEPC and COCA general corpus -Learners' report of findings on self-constructed mini business corpus -Learners' report of findings on Learners Corpus by use of AntConc

#### 5.4.4 Post-test

After 6-week sessions, the instructed learners took part in the post-test of collocation competence, which was in

the same form of the pre-test. And an online survey, in a Likert scale, was administered to see into learners' recognition of changes of collocation awareness, learner autonomy and collocation competence.

## 6. Data Analysis and Discussion

### 6.1 Research Question 1: Collocation Awareness

Relevant data were generated and collected from the questionnaire and learners' reflective journals.

#### 6.1.1 Questionnaire Data

The data about learners' collocation awareness were collected from four multiple-choice-questions in the format of Likert Scale measures respondents' recognition of their collocation awareness. The four question titles are: 1. Before the experiment, I attached importance to business phrases (especially collocations); 2. Since the experiment, I have paid attention to collocational structures in business contexts in Comprehensive Business English Course (CBEC); 3. Since the experiment, I have attention to collocational structures out of CBEC; 4. Since the experiment, I have practiced the collocational structures learned intentionally. The five alternative choices are "never", "seldom", "sometimes", "often" and "always".

Table 4 indicates that before the experiment, the learners had comparatively weak collocation awareness in that only 21.7% often paid attention to collocations in business context, 43.5% of them sometimes did so, and all the remaining 34.8% seldom noticed them.

Table 4. Learners' collocation awareness before the experiment

Scale	Frequency	Percent	Cumulative Percent
always	0	0	0
often	5	21.7	21.7
sometimes	10	43.5	65.2
seldom	8	34.8	100.0
never	0	0	100.0
Total	23	100.0	

From Table 4 and Table 5, it can be seen that the percentage figure of the learners who seldom noticed collocations drops from 34.8% to 4.3% and over half of the learners stated that they often paid attention to collocations in Comprehensive Business English Course since the experiment.

Table 5. Learners' collocation awareness in CBEC since the experiment

Scale	Frequency	Percent	Cumulative Percent
always	1	4.3	4.3
often	12	52.2	56.5
sometimes	9	39.2	95.7
seldom	1	4.3	100.0
never	0	0	100.0
Total	23	100.0	100.0

And the collocation awareness out of this course increased accordingly as indicated in Table 6. All the learners stated that they also attached importance to collocations out of Comprehensive Business English Course, with 56.5% of them sometime did so and 43.5% often did so.



Table 6. Learners' collocation awareness out of CBES since the experiment

Scale	Frequency	Percent	Cumulative Percent
always	0	0	0
often	10	43.5	43.5
sometimes	13	56.5	100.0
seldom	0	0	100.0
never	0	0	100.0
Total	23	100.0	100.0

Collocation awareness in this article is defined as “intentional noticing, understanding and practicing collocational meaning of key words”; therefore, whether the learners take the initiative in practicing those noticed collocations is equally important as noticing them. Table 7 demonstrates that since the experiment, over 90% learners stated that they sometimes or often put those noticed and understood collocations into use in or out of class.

Table 7. Learners' awareness of practicing collocations since the experiment

Scale	Frequency	Percent	Cumulative Percent
always	0	0	0
often	3	13.0	13.0
sometimes	18	78.3	91.3
seldom	2	8.7	100.0
never	0	0	100.0
Total	23	100.0	100.0

The paired T-test results presented in Table 8 reinforced the preliminary findings from the above descriptive data.

Table 8. Paired samples test

Pairs				Paired Differences				
				Mean	Std. Deviation	t	df	p (2-tailed)
Pair 1	Collocation awareness before the experiment	awareness of noticing collocations in CBEC	before the experiment	-.696	.703	-4.746	22	.000
Pair 2	Collocation awareness before the experiment	awareness of noticing collocations out of CBEC	before the experiment	-.565	.788	-3.441	22	.002
Pair 3	Collocation awareness before the experiment	awareness of practicing collocations since the experiment	before the experiment	-.174	.887	-.940	22	.357

p<.05.

The above data showed that learners' collocation awareness from the perspective of reception improves significantly from the input perspective (Pair 1 and 2) while no significant change has been seen from the production (Pair 3). That is to say, learners' initiative efforts in noticing and understanding collocations have increased significantly while that in producing them not.

The framework for investigating L2 acquisition developed by Ellis (1994) in Figure 7 can account for this

difference. It specifies how L2 acquisition begins from noticed input to L2 output. When input is noticed and interpreted at different stages of development and organized as different representations of knowledge in order to be stored for L2 output, “intake” and “two representations of explicit knowledge and implicit knowledge (interlanguage system)” are two important stages that cannot be skipped and really take time. In this sense, even though significant increase of collocation awareness in the input stages can be detected, no immediate significant increase in the output stage can be found accordingly. But the teacher and teaching materials designer can make better use of the corpus data to facilitate the development of this process in that more output practices could be added into the in-class activities.

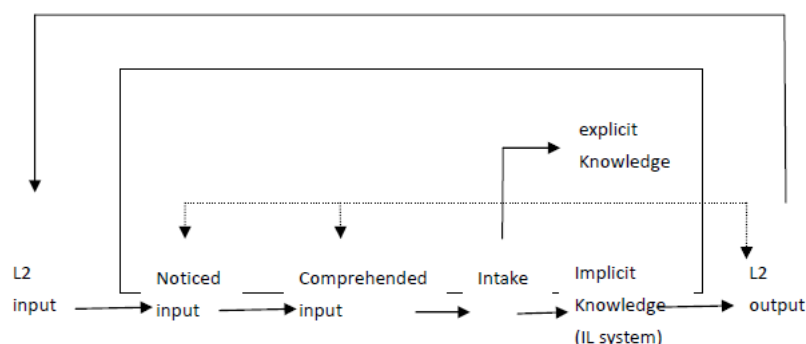


Figure 7. A framework for investigating L2 acquisition

Even though, Schmit (1990) argued that noticing is the necessary and sufficient condition for converting input to intake when he defined the role of consciousness in Second Language Learning. Rutherford and Smith (1985) called this kind of explicitly drawing learners' attention to the formal properties of language in order to facilitate language learning as “consciousness-raising”. And consciousness is in both usage and theoretical treatment of the topic is commonly equated with awareness. A central element of language teaching is raising students' awareness of, and developing their ability to ‘chunk’ language successfully (Lewis, 2012). Among different groups of chunks, the most important two are collocations, which are message-oriented, and institutionalized expressions, which are essentially pragmatic in character (Lewis, 2012). Lewis suggests that teachers should raise students' awareness of collocations and initiate their own action research to make sure the changes they make are of benefit to their students.

#### 6.1.3 Quotes from Reflective Journals

Qualitative data from learners' reflective journals also support the above findings from quantitative data. As in an open form, learners' responses varied, but still some recurrent ones could be extracted to indicate the increase of learners' collocation awareness.

I can know what words are always together with the word I check, which helps me get a better command of the use of the word.(Student a, Journal 1)

I need not only to know the words, but know how to use them and how to combine the words. And the application itself is a good way to help me to remember them, from the usage. (Student b, Journal 1)

Simultaneously, right collocation is very important in Business English, so I am supposed to pay more attention to it. For example, I must figure out which words are correct to modify the word “economy” in Business English.(Student c, Journal 1)

Because I need to accumulate the set structures of the phrases and fixed usages of the words in the business field. (Student d, Journal 1)

I discover that I still need to enhance my vocabulary concerning to business and finance, as I sometimes need to figure out the different meaning of some familiar collocations, and the difference can be significant. (Student e, Journal 2)

I also agree on the idea that we should lay our emphasis on phrases and collocations. This is an efficient and vivid way to learn language skills. (Student f, Journal 2)

But corpus gives us a chance to search the real using frequencies of collocations about those words. It shows the

field that these words used in at the same time, which is good for us to understand those words better than ordinary dictionary. (Student f, Journal 2)

Those extracted lines from learners' reflective journals collectively communicate that:

- 1) Collocations are a very important language feature to notice and master in L2 Business English acquisition;
- 2) They had relatively low awareness of collocation in Business English learning in the past;
- 3) They have realized the importance of acquisition of collocations in business and begun to do it.

Those recurrent ideas from learners' journals during the experiment perfectly indicate a raised collocation awareness of learners.

## 6.2 Research Question 2: Learner Autonomy

Boosted learner autonomy seems equally important as increased collocation awareness when it comes to collocation learning in that only with lasting autonomous learning can learners conduct sustainable language learning with the aid of corpus tools. The present section is aimed to address the research question whether learners have stepped closer to an autonomous learner by use of corpus tools. The questions in Part 2 of the questionnaire, also in Likert scale, specify this question through four variables of learner autonomy of corpus tools: 1) mastery of corpus tools; 2) frequency of corpus tools use; 3) frequency of exchanging ideas about usage of corpus tools; 4) willingness to use corpus tools for self-access learning in the future. The five alternative choices in the first question of the first question are "very poor", "poor", "fair", "good" and "very good". Those for the other three questions are "never", "seldom", "sometimes", "often" and "always".

### 6.2.1 Mastery of Corpus Tools

As indicated in Table 9, 39.1% learners claimed to have good mastery of the corpus tools and skills and 47.8% stated that they did a fair job, with only 13% learners thought they had a poor mastery.

Table 9. Mastery of corpus tools

Scale	Frequency	Percent	Cumulative Percent
very good	0	0	0
good	9	39.1	39.1
fair	11	47.8	86.9
poor	3	13.1	100.0
very poor	0	0	100.0
Total	23	100.0	100.0

An interview was conducted on specified respondents of the questionnaire, who gave negative feedbacks on mastery of corpus tools, and it has been found that the relatively poor mastery of corpus tools was attributed to two major reasons. Firstly, inadequate instructions on COCA and insufficient practice in class, and wordy reader's book on COCA instruction for after-class reading made it less effective for undergraduate learners to master COCA as a self-access learning tool. Secondly, COCA's daily limit on user's queries to a great extent constrained learners' exploration.

### 6.2.2 Frequency of Corpus Tools Use

Despite the less incomplete mastery of corpus tools, the instructed learners demonstrated their frequent practice of corpus-aided self-access learning in and out of this course. Table 10 shows that all the learners reported use of corpus tools in Comprehensive Business English Course, with 4.3% of them always and 47.8% sometimes did so. In another two ongoing language-focused courses Business English Listening and Business English Writing, they also conducted corpus-aided self-access learning. What's worth noticing is that 73.9% learners stated that they always, often or sometimes used corpus tools in Business English Writing Course, which indicates that corpus tools were more often used as a self-access learning tool to facilitate writing. Apart from those language-focused courses, some learners also reported a list of business-related courses, in which they conducted corpus-aided learning, like Principle of Management, Business Communication, Principle of Accounting, Principles of Economics, and Marketing. What's more encouraging is that some learners even used corpus tools for out-of-class journal reading, preparation for speeches, watching English dramas and online job application,

etc. Corpus tools introduced in Comprehensive Business English Course have been frequently and widely used among the instructed learners.

Table 10. Use of corpus tools in language-focused courses

Scale	Comprehensive English		Business			
	Frequency	Percent	Business English Listening		Business English Writing	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
always	1	4.3	0	0	2	8.7
often	3	13.0	4	17.4	6	26.1
sometimes	11	47.9	6	26.1	9	39.1
seldom	8	34.8	9	39.1	4	17.4
never	0	0	4	17.4	2	8.7
Total	23	100.0	23	100.0	23	100.0

### 6.2.3 Frequency of Exchanging Ideas about Corpus-aided Learning

During the teaching experiment, 65.2% learners stated that they sometimes or always exchanged ideas about corpus-aided learning. Only 1 student demonstrated that they had never talked about corpus-aided learning with peers and the instructor.

Table 11. Frequency of Exchanging Ideas about Corpus-aided Learning

Scale	Frequency	Percent	Cumulative Percent
always	0	0	0
often	3	13.0	13.0
sometimes	12	52.2	65.2
seldom	7	30.4	95.7
never	1	4.3	100.0
Total	23	100.0	

### 6.2.4 Willingness to Use Corpus Tools for Self-access Learning in the Future

The writer holds that learner's autonomy can be best demonstrated in their sustained corpus-aided self-access learning. According to the data from the questionnaire, nearly all the respondents gave a positive reply to whether to conduct corpus-aided learning in the future, with 100% of them stating willingness in language-focused courses and 95.7% of them in business-related courses. About 87% of the respondents reported they would sometime, often or always use corpus tools to conduct self-access learning in language-focused courses.

Table 12. Willingness to conduct corpus-aided self-access learning in the future

Scale	Language-focused courses			Business-related courses		
	Frequency	percent	Frequency	percent	Frequency	percent
always	1	4.3	4.3	0	0	0
often	6	26.1	30.4	12	52.2	52.2
sometimes	13	56.5	86.9	8	34.8	87.0
seldom	3	13.0	100.0	2	8.7	95.7
never	0	0	0	1	4.3	100.0

Scale	Language-focused courses		Business-related courses			
	Frequency	percent	Frequency	percent	Frequency	percent
always	1	4.3	4.3	0	0	0
often	6	26.1	30.4	12	52.2	52.2
sometimes	13	56.5	86.9	8	34.8	87.0
seldom	3	13.0	100.0	2	8.7	95.7
never	0	0	0	1	4.3	100.0
Total	23	100.0	23			

When interviewed about why he is reluctant to follow corpus-aided approach for self-access learning, the only learner who expresses unwillingness regarded it as time-and-labor consuming compared to consulting a dictionary.

Based on the data from the four variables, it has been found that: 1) The instructed learners have generally developed the skills required for corpus-aided self-access learning; 2) Despite some learners' less immediate mastery of corpus skills, the majority adopted corpus-aided self-access learning in both language-focused courses and business-related ones, with the highest frequency in Business English Writing Course; 3) Most learners exchanged ideas about corpus techniques during the experiment; and 4) Nearly all the learners expressed willingness to adopt corpus-aided self-access learning sustainably. These findings brought to a natural conclusion that the instructed learners have been inspired, instead of "taught" to be an autonomous learner in business and English study. The mechanism could be strongly supported by Boulton (2009) as follows.

In terms of autonomy, corpus-aided pedagogy places responsibility on learners to discover patterns in the data presented. It may even be, in some cases, that learning is more effective without a teacher, i.e. when learners discover things for themselves under the guidance of teacher after mastery of essential tools. This is the basic premise of data-driven learning, where learners examine naturally-occurring language and discover patterns on their own. It is alleged to have many advantages- to foster learner autonomy, increase language awareness, and improve ability to deal with authentic language, and so on.

### 6.3 Research Question 3: Collocation Competence in a Short Term

#### 6.3.1 Data from Pre-test and Post-test

Table 13 shows that the mean score of the instructed group rose from 13.8 in the pre-test to 18.0, with an increase of 30%. The paired T-test proves a statistically significant difference between the pre-test and post-test ( $p=.000<.05$ ) That suggests a significant rise in the instructed learners' collocation performance in the test, which in turn implies a significant improvement of collocation competence since the conduct of corpus-aided pedagogy.

Table 13. T-test of Collocation Competence Tests (two-tailed)

	N	Mean	Standard Deviation	t	p
Pre test	23	13.8	3.704	-4.834	.000
Post-test		18.0	3.723		

$p<.05$ .

#### 6.3.2 Data from Questionnaire

Data from questionnaire strongly support the findings from the pre-test and post-test. When asked whether they acknowledge a significant improvement in recognizing, comprehending and using the key collocations in Business English since the corpus-aided pedagogy, 69.6% instructed learners expressed their content with a tick on "mostly agree".

Table 14. Learners' Acknowledgement of a Significant Increase in Collocation Competence

Scale	Frequency	Percent	Cumulative Percent
completely agree	0	0	0
mostly agree	16	69.6	69.6
don't know	5	21.7	91.3
mostly disagree	2	8.7	100
completely disagree	0	0	100
Total	23	100	100

### 6.3.3 Data from Learners' Reflective Journals

From the learners' reflective journals, the relevant reasons for this significant improvement of collocation competence could be generalized as follows.

- 1) Compared to traditional grammar-translation pedagogy, the corpus-aided pedagogy in theme-based business English instruction gives learners a closer look at the collocation structures in certain business topics, which shifts their focus onto key business collocation in business context, which is the first step in improvement of collocation competence;
- 2) Most learners take their initiative in self-access learning with the aid of AntConc and COCA with the focus on collocations of key words, which boosts learner autonomy. During this self-access learning, they have to discover by clues from the corpus-aided data instead of being told, which is crucial for improvement of collocation competence;
- 3) Approach a large Business English corpus by use of AntConc gives learners full exposure to abundant natural and idiomatic language resources than a small number of textbook reading passages, which facilitates improvement of collocation competence.

Data from the tests, the questionnaire and learners' reflective journals hold collaboratively that, exposed to rich corpus resources, Business English learners in corpus-aided pedagogy can raise their collocation awareness and competence in a relatively short period of time, with the rigorous support of corpus-aided self-access learning by use of corpus tools and under proper guidance of the instructor.

### 6.3.4 Learners' Proficiency Level and Improvement in Learners' Collocation Competence

In order to gain a deeper insight into the intermediary role of language proficiency in the interaction between corpus-aided pedagogy and learners' collocation competence, the instructed learners were divided into three subgroups—High Level, Intermediate Level and Low Level with proportion of 30%, 40% and 30% respectively. The T-test of the three subgroups indicates that High Level and Intermediate Level subgroups have witnessed a significant improvement in their collocation competence (High Level:  $p=.016<.05$ ; Intermediate Level:  $p=.015<.05$ ) while the increase in Low level subgroup has not been proven to be statistically significant ( $p=.080>.050$ ).

Table 15. Paired T-Test for grouped learners

	High Level	Intermediate Level	Low Level
N	7	9	7
Pre-test mean (standard deviation)	15.3(2.752)	14.0(4.950)	12.0(1.915)
Post-test mean (standard deviation)	19.6(3.505)	17.8(3.346)	16.6(4.276)
t	-3.333	-3.091	-2.107
p	.016	.015	.080

$p<.05$ .

After a close look at the responses of the seven learners in Low Level subgroup about their mastery of corpus tools and intentional use of noticed collocations, it has been found that though they sometimes or always used

the intentionally noticed collocations, five of them didn't think they had a good mastery of the techniques of the corpus tools. The feedback in their reflective journals seem to bear out this claim, as the following quotes demonstrate their ideas that corpus tools are not effective for self-access learning in that it is labor-consuming and time-consuming though collocation-focused Business English instruction is interesting.

As a matter of fact at first I thought that these full English corpuses were too obscure to use and I get used to using Youdao Dictionary or Kingsoft PowerWord to check English words. However, when I try to use COCA, I find that this kind of corpus does more good to my English study. I can read a number of sentences so that I can remember English words with the help of different contexts. I can know what words are always together with the word I check, which helps me get a better command of the use of the word. COCA is really good. Though it is not easy to operate, I can make more attempts in order to master it. (Student g, Journal 1)

I think we have gotten used to some online dictionaries because of their convenience. Because the method of using corpus is a bit complex, I don't consider it good to use the corpus to check out unknown words at usual. (Student h, Journal 2)

In the class on Monday, some of my classmates showed us how to use COCA and PIGAI. Although I have known the two useful websites that can help me in English writing and solve several vocabulary problems on the websites, my classmates taught me more. They taught us the convenient way to find synonym and right collocation; they also showed us a specific process of correcting an article completely. I learn a lot from their presentation, especially the use of COCA. I used to differentiate similar words on COCA, while after the class; I know more skills of using corpus. ... Now, I only use AntConc once. I think it is a little bit difficult for a beginner like me. Maybe I can be proficient in using AntConc after several times. (Student i, Journal 2)

You may know the original meaning of the word, but you may don't know the meaning in marketing... This week I used "COCA" to check several words, and I found it not so convenient for me, and I have to search what I want to talk about for long time because there are too many options. However, the collocation is surely correct and accurate; when I found it, it should be a perfect answer which I don't have to be nervous about whether it's correct. For example, I want to say something but don't know how to say it in English, it can't help me, and I have to find the specific words and their collocations to be complete sentences. It's really annoyed me. (Student j, Journal 2)

From the quoted reflective scripts, it can be told that learners in the Lower Level subgroup regarded use of corpus tools as complex and laborious and thus reported a less effective mastery of them although they acknowledged the benefits of corpus-aided collocation pedagogy. This maybe accounted for their statistically insignificant improvement in the tests. Without the effective aid of techniques and resistance of Observe-Hypothesis-Experiment cycle, they couldn't benefit from corpus-aided self-access learning in a real sense.

## 7. Conclusions

All the above discussions on collected data from tests, questionnaire and learning reflective journals have addressed the three research questions adequately.

It has been found that after the 8 weeks' teaching experiment, the instructed learners' collocation awareness in regard of noticing and comprehending has been significantly raised, while no significant increase of their intentional use of collocations has been reported probably because of the gap between input and output in L2 acquisition. Even so, corpus-aided awareness-raising activities have been regarded as effective by learners in that they awakened their consciousness of collocations in Business English.

With enhanced awareness in collocations, EFL learners, guided by the teaching ideology of Observe-Hypothesis-Discovery, have been motivated to be a more autonomous learner by use of corpus tools. Teacher's role in this process shifted from a source of information to an autonomous learning facilitator.

Although the instructed learners didn't report a significant increase in their intentional use of collocations, the tests results still revealed a significant improvement of collocation competence in a short term, which corresponded to data from the questionnaire about their recognition of improvement of collocation competence. In a short term, the instructed learners have gained a significant improvement in their performance of collocation test of selected key words, with the learners with high and intermediate English proficiency benefit more than those with low English proficiency, which may indicate corpus-aided Business English pedagogy is more effective for intermediate and advanced English learners.

## 8. Limitations

One major limitation of the present study is that the time frame of the teaching experiment is not enough to allow the effect of corpus-aided pedagogy to be fully revealed, but it does suggest that the potential long-term advantage, i.e. collocation awareness and learner autonomy outweigh the short-term outcome in collocation competence examined in the present research. Besides, the corpus tools introduced in the empirical study could be diversified for different specific purposes in the teaching process in the future study.

## Acknowledgments

This paper is funded by the 2015 Higher Education Teaching Research and Reform Project of Guangdong Province “Corpus-aided Studies of Business English Phrases Pedagogy “ and 2014 Joint Research Program of Guangdong University of Foreign Studies “Two Dimensional Analysis of Chinese Learners’ Use of Shell Nouns: A Corpus-based Study”.

## References

- Allan, R. (2009). Can a graded reader corpus provide “authentic” input? *ELT Journal*, 63(1), 23-32. <https://doi.org/10.1093/elt/ccn011>
- Boulton, A. (2009). Testing the limits of data-driven learning: language proficiency and training. *ReCALL*, 21(1), 37-54. <https://doi.org/10.1017/S0958344009000068>
- Chen, L. D. (2015). A corpus-based study of shell nouns in “N + that” appositive clause in Business English undergraduate theses. *Journal of Zhejiang International Studies University*, 6, 28-35. <http://doi.org/10.3969/j.issn.2095-2074.2015.06.005>
- Du, H. Z. (2013). Study of corpus-based oral Business English teaching approach. *Journal of Language and Literature Studies*, 3, 124-126.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O’Keeffe, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359-371). London: Routledge. <https://doi.org/10.4324/9780203856949.ch26>
- Hu, C. Y. (2011). Corpora and the teaching of collocation in Business English. *Journal of Shaoguan University*, 32(5), 205-208.
- Hu, C. Y. (2012). Corpus and Business English pedagogy. *English Teachers*, 6, 42-47.
- Johns. T. (1991). Should you be persuaded: two samples of data-driven learning materials. *ELR Journal*, 4, 1-16.
- Lewis, M. (2000). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation: further developments in the lexical approach*. England: Language Teaching Publications.
- Lewis, M. (2012). *The Lexical Approach: The State of ELT and a Way Forward*. Hampshire, UK, Cengage Learning CMEA.
- Lin, L. (2016). Application of business contract parallel corpus in Business English translation pedagogy. *Science & Technology Information*, 25, 138-139.
- Nelson, M. (2000) *A corpus-based study of Business English and Business English teaching materials*. (Unpublished doctoral dissertation). Manchester: University of Manchester.
- Nelson, M. (2006). Semantic associations in Business English: A corpus-based analysis. *English for Specific Purposes*, 25, 217-234. <https://doi.org/10.1016/j.esp.2005.02.008>
- Poole, R. (2016) A corpus-aided approach for the teaching and learning of rhetoric in an undergraduate composition course for L2 writers, *Journal of English for Academic Purposes*, 21, 99-109. <https://doi.org/10.1016/j.jeap.2015.12.003>
- Rutherford, W., & Sharwood, S. (1985). Consciousness-raising and universal grammar. *Applied Linguistics*, 6, 274-82. <http://dx.doi.org/10.1093/applin/6.3.274>
- Schmit, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 17-46. <https://doi.org/10.1093/applin/11.2.129>
- Shen, T. (2016). An empirical study of teaching chunks in oral Business English from the perspective of Schema Theory. *Shanxi Education*, 3, 37-39.
- Shi, J. (2012). Application of Sketch Engine in Business English lexis pedagogy. *English Teachers*, 12(7), 53-60.



- Sun, H. Y. (2008). An analysis of the development of China's EFL learners' collocational competence. *Foreign Languages Research*, 108(2), 56-71. [http:// doi.org/10.13978/j.cnki.wyyj.2008.02.019](http://doi.org/10.13978/j.cnki.wyyj.2008.02.019)
- Walker, C. (2011). How a corpus-based study of the factors which influence collocation can help in the teaching of business English. *English for Specific Purposes*, 30(30), 101-112. <https://doi.org/10.1016/j.esp.2010.12.003>
- Wang, D. M. (2009). On corpus-based Business English vocabulary teaching. *Journal of Suzhou Education Institute*, 12(2), 143-145. [http:// doi.org/10.13985/j.cnki.34-1227/c.2009.02.032](http://doi.org/10.13985/j.cnki.34-1227/c.2009.02.032)
- Wang, L. F., & Huang, X. Q. (2011). Developing university computer-aided Business English writing teaching system. *Technology Enhanced Foreign Language Education*, 142(11), 37-41.
- Wei, H., Zhang, P., & Zhong, H. (2014). Business English translation pedagogy based on computer-based corpus. *Journal of Fujian Institute of Education*, 15(4), 103-106.
- Zhang, L. Y. (2012). Online corpus-based corpus Business English writing model studies. *Journal of Language and Literature Studies*, 5, 161-163.

### Notes

Note 1. From the perspective of semantic prosody, Nelson, based on his research of Business English Corpus, categorized business lexis into five patterns: people in business, business activities, business description, business events and entities. Based on this, the pre-test revolves around 10 business keywords from the pedagogical corpus: manger, employee, business, market, describe, access, raise, financial, sale and ownership.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).