Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 2, June 2025 Pp.233-261
DOI: https://dx.doi.org/10.24093/awej/vol16no2.14

Examining the Effectiveness of Sketch Engine for Verb-noun Collocation Learning among Chinese Low-Proficiency Learners

Shuang Yang
School of Education, Faculty of Social Sciences and Humanities,
Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100, Kuala Lumpur, Malaysia
Corresponding Author: yangshuang@graduate.utm.my

Wan Nur Asyura Wan Adnan
School of Education, Faculty of Social Sciences and Humanities,
Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100, Kuala Lumpur, Malaysia
Email: asyura.kl@utm.my

Received: 02/25/2025 Accepted:05/28/2025 Published:06/26/2025

Abstract
Developing collocational competence is essential for Second Language proficiency, yet low-proficiency learners often struggle with verb-noun collocations due to the arbitrary nature of collocational patterns. While influential corpus tools such as the Corpus of Contemporary American English and the British National Corpus serve as valuable resources for collocation learning, they also pose significant cognitive challenges for low-proficiency learners. Therefore, this study investigated the effectiveness of corpus tools integrated with pedagogical practices in facilitating verb-noun collocation learning among Chinese low-proficiency learners, to identify the most suitable tool tailored to their specific learning needs. A quantitative research method was employed, utilizing an intervention design with pre-and post-tests to compare the impact of using Sketch Engine for Language Learning (SkELL) and COCA in enhancing learners' verb-noun collocation proficiency. Furthermore, the eight-week intervention integrated SkELL and COCA into collocation lessons respectively, each designed within the framework of Corpus-Based Language Pedagogy (CBLP), presenting how strategic instruction maximized the potential of corpus tools in facilitating learners' improvement in verb-noun collocation proficiency. The findings of the study revealed that SkELL significantly improved verb-noun collocation proficiency among Chinese low-proficiency EFL learners, attributed to its intuitive interface, simplified search functions, and scaffolded learning environment. Moreover, this study offers valuable contributions to the design of corpus-based instruction by demonstrating the effectiveness of SkELL as a corpus tool tailored to the needs of Chinese low-proficiency EFL learners. The study concluded with insights on pedagogical implications and recommendations for further research.

Keywords: Corpus-Based Language Pedagogy, Corpus of Contemporary American English, Low-proficiency EFL learners, Sketch Engine for Language Learning, verb-noun collocation
How to cite: Yang, S.& Adnan, W.N.A. (2025). Examining the Effectiveness of Sketch Engine for Verb-noun Collocation Learning among Chinese Low-Proficiency Learners. *Arab World English Journal, 16*(2): 233-261. https://dx.doi.org/10.24093/awej/vol16no2.14

Introduction

Collocational competence is a crucial aspect of Second Language (L2) proficiency, particularly in academic and professional communication. Among various types of collocations, verb-noun collocations, with highly frequent use, play a central role in shaping fluency and accuracy in L2 writing and speaking. However, research indicates that L2 learners, especially low-proficiency learners, frequently struggle with acquiring and using verb-noun collocations correctly due to the opaque and arbitrary nature of collocational patterns (Wray, 2002).

To address these challenges, corpus tools have been increasingly integrated into collocation instruction, providing learners with access to authentic language use and empirical linguistic patterns (Boulton & Cobb, 2017). Tools such as the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC) have been widely adopted to support collocation learning, particularly within data-driven learning (DDL) field (Akıncı & Yıldız, 2017; Al-Mahbashi et al., 2015; Altun, 2021; Daskalovska, 2015; Durmuş & Kilimci, 2020; Girgin, 2019; Jafarpour et al., 2013; Li, 2017; Łuszcz, 2016; Ma & Mei, 2021; Men, 2020; Satake, 2023; Ucar & Yükselir, 2015; Wu, 2021). However, the complexity of their interfaces, coupled with intricate data retrieval processes, may hinder learners' ability to effectively analyse and apply collocational patterns. While research consistently demonstrates that corpus tools enhance collocational proficiency, findings suggest that their effectiveness is often more pronounced among higher-proficiency learners than among low-proficiency learners (Boulton & Cobb, 2017; Daskalovska, 2015; Wu, 2015). Such debate raises concerns about whether corpus-based learning is equally beneficial for all learner groups and how instructional interventions can be designed to support low-proficiency learners in leveraging corpus resources effectively.

Undeniably, while exposure to authentic texts that naturally incorporate collocations enhances learners' understanding, low-proficiency English learners may struggle with acquiring collocations. The overwhelming amount of fragmented sentences presented in a short period can make the learning process more challenging for them (Boers et al., 2014; Daskalovska, 2015). Recently, a new corpus tool for English language learning, known as Sketch Engine for Language Learning (SkELL), has been introduced as an alternative resource. Unlike the previously mentioned corpora, SkELL provides a sufficient number of complete sentences, enabling students to learn multiple patterns in context. Regrettably, little research has been conducted to examine its effectiveness, compared with COCA as a long-influential tool in China, on collocation knowledge among low-proficiency learners.

There is a growing trend toward integrating corpus tools into collocation learning, with promising yet varied results across proficiency levels. Despite advancements in corpus-based instruction, gaps remain in understanding how different tools influence collocation acquisition among low-proficiency learners. Previous studies have largely focused on COCA and BNC (Akıncı & Yıldız, 2017; Al-Mahbashi et al., 2015; Altun, 2021; Ayyat, 2024; Daskalovska, 2015; Durmuş & Kilimci, 2020; Girgin, 2019; Jafarpour et al., 2013; Li, 2017; Łuszcz, 2016; Ma & Mei, 2021; Men, 2020; Satake, 2023; Ucar & Yükselir, 2015; Wu, 2021), with limited exploration of simplified alternative tool like SkELL. Furthermore, while the effectiveness of corpus tools in general collocation learning has been established, research is still lacking on how different corpus interfaces, training approaches, and scaffolding techniques impact verb-noun collocation learning specifically among struggling learners.

Given the indispensable role of appropriate corpus tools, especially for lower-proficiency learners, in verb-noun collocation learning, the strategic pedagogy that uses corpora and corpus tools in DDL can help low-proficiency learners avoid confusion and cognitive overload. Consequently, Ma et al. (2022) introduce the concept of Corpus-Based Language Pedagogy (CBLP), which emphasizes the integration of corpus technology into language learning through a strategically designed curriculum and instructional strategies. For example, a teacher prioritizes selecting an appropriate corpus tool tailored to low-proficiency students, providing supplementary training, and designing a structured corpus-based collocation lesson to enhance their learning experience.

Notably, no prior research has applied SkELL in collocation learning for Chinese EFL low-proficiency students. Existing studies have predominantly explored SkELL's effects on Japanese learners, leaving a significant gap in understanding how Chinese EFL learners interact with this tool. Given the linguistic and educational differences between Japanese and Chinese learners, our study addresses this critical gap by examining how SkELL-supported collocation instruction affects Chinese students' verb-noun collocation learning compared to COCA. The study not only expands the applicability of SkELL but also contributes to a broader discussion on corpus tool suitability across different learner populations.

To enhance verb-noun collocation proficiency among low-proficiency learners, the study aims to examine the effectiveness of corpus tools integrated with pedagogical practices and to identify the most appropriate tool tailored to their learning needs. In the field of corpus linguistics, this study makes a significant contribution by informing the corpus-based lesson design and the selection of corpus tools specifically tailored to meet the learning needs of Chinese low-proficiency learners. Building upon these insights, the study implemented an intervention design featuring an eight-week instructional lessons grounded in CBLP. The intervention incorporated structured lesson plans, scaffolded activities, and guided exploration of corpus tools, ensuring accessibility and effectiveness for low-proficiency learners.

To shed more light on the above study objectives, an intervention was conducted to compare the impact of SkELL and COCA in enhancing English verb-noun collocation learning among Chinese EFL low-proficiency learners by addressing the following research questions:

1. To what extent do SkELL and COCA enhance verb-noun collocation proficiency among low-proficiency Chinese EFL learners?
2. How do SkELL- and COCA-supported lessons differ in their impact on verb-noun collocation acquisition for low-proficiency Chinese EFL learners?
3. How do learners perceive the effectiveness of SkELL for verb-noun collocation learning?

Literature Review

Social Constructivism Theory (SCT) in Designing Corpus-based Learning

Constructivist theories of learning emphasize the active role of learners in constructing knowledge through interaction with their environment. In the context of Second Language (L2) collocation acquisition, constructivism highlights the importance of meaningful engagement, scaffolded learning, and the co-construction of knowledge through guided discovery (Vygotsky, 1978), while it offers a foundational framework for understanding how learners develop linguistic competence through mediated learning experiences. Central to this theory is the concept of the

Zone of Proximal Development (ZPD), which describes the difference between what a learner can achieve independently and what they can accomplish with the assistance of a more knowledgeable other. ZPD underscores the necessity of scaffolding, where structured support is gradually removed as learners gain proficiency (Tabib, 2022, p. 187). In language learning, scaffolding can take the form of teacher guidance, peer collaboration, and technological tools that facilitate corpus exploration.

Corpus-based Language Pedagogy (CBLP) aligns closely with Vygotsky's SCT theory, as it provides learners with authentic language input and allows them to interact with linguistic data through data-based learning. Within a constructivist paradigm, corpus tools such as COCA and SkELL serve as mediating artefacts that bridge the gap between learners' existing knowledge and their potential development in collocation competence. For low-proficiency learners, direct engagement with large corpus databases may present cognitive challenges, as observed in previous studies (Boers et al., 2014; Boulton & Cobb, 2017; Daskalovska, 2015; Liu & Ma, 2025). However, when combined with structured scaffolding, corpus tools become effective learning instruments. Teachers act as facilitators, guiding students in interpreting corpus output, identifying collocational patterns, and applying them in their target tasks. This mediated learning process enables students to move through their ZPD and develop collocational proficiency in a way that would not be possible through traditional memorization-based approaches. Hence, a crucial way for effective corpus-based learning is selecting tools that cater to low-proficiency learners' linguistic abilities. SkELL, for instance, offers a simplified interface that reduces cognitive load, making it more accessible for low-proficiency learners. By integrating corpus tools into guided instructional activities, teachers provide the necessary scaffolding to help learners analyze and apply verb-noun collocations meaningfully.

By situating this study within the SCT framework, this study emphasizes the interactive and scaffolded nature of corpus-based collocation learning. In particular, Vygotsky's notion of the ZPD highlights the importance of structured guidance in helping low-proficiency learners navigate corpus tools effectively. Furthermore, the concept of CBLP reinforces the idea that successful corpus-based instruction requires both appropriate tool selection and well-designed scaffolding strategies (Liu & Ma, 2025), which informs the present study's design, examining the role of both corpus tools and instructional scaffolding in enhancing verb-noun collocation proficiency for low-proficiency learners.

Corpus Tools for Verb-Noun Collocation Learning

The application of corpus tools for verb-noun collocations has gained increasing scholarly attention in the last two decades. Early corpus-based studies primarily focus on the benefits of data-driven learning (DDL), where learners interact with authentic language examples extracted from large corpora (Boulton & Cobb, 2017; Crosthwaite, 2024). DDL approach revalues collocation teaching by focusing on learner discovery rather than traditional rule-based methods. Nowadays, the focus of research has shifted from determining whether DDL works to exploring how it can be effectively implemented for diverse language proficiency learners (Alharbi, 2024). Notably, tools in DDL such as the British National Corpus (BNC) and Corpus of Contemporary American English (COCA) have been widely used to support collocation learning by providing contextualized instances of word combinations. Teaching collocations through corpus-based tools

was of significant help to the students' retention and learning of collocations (Akbari et al., 2015). Many studies underscore the need for DDL with explicit instruction in collocation instruction (Akıncı & Yıldız, 2017; Frankenberg-Garcia, 2024; Gablasova et al., 2017; Liu & Gablasova, 2023), while others highlight the effectiveness of corpus tools like Sketch Engine in enabling learners to identify recurring patterns in authentic texts (Hirata & Hirata, 2019; Kilgarriff et al., 2015). For instance, Chen (2017) conducted an empirical study on business English learners and found significant gains in collocational competence following an eight-week intervention. Sun and Park (2023) highlight the effectiveness of corpus-based collocation instruction in improving learners' collocation knowledge and retention, while also noting the need to resolve conflicting results regarding instructional methods. These findings indicate that corpus tools play a crucial role in collocation acquisition but may require tailored pedagogical approaches for different-level students to maximize benefits.

Consistent with these claims, recent studies argue that with appropriate scaffolding in DDL, low-proficiency learners can benefit significantly from corpus tools. Chen (2017) and (Gablasova et al., 2017) found that structured corpus-based instruction led to measurable improvements in collocational accuracy, even among less proficient learners. They tend to support that corpus-based pedagogy is proved to be effective for lower-level learners. The introduction of user-friendly corpus tools, such as SkELL (Sketch Engine for Language Learning), has been a pivotal development in addressing cognitive barriers by providing more accessible search functionalities and intuitive collocation outputs (Hirata & Hirata, 2019; Kilgarriff et al., 2015). As they elaborate, SkELL is a valuable educational tool that broadens students' English learning experiences through authentic examples and offers opportunities for teachers to create more communicative activities, though further research is needed for its effective implementation.

Subsequently, one of the critical debates in corpus-based collocation learning is its effectiveness across different proficiency levels. While research consistently supports the effectiveness of corpus tools for advanced learners, findings regarding low-proficiency learners remain inconsistent (Boulton & Cobb, 2017; Daskalovska, 2015; Wu, 2015; Yoon, 2008). For example, studies examining the effectiveness of corpus-based collocation pedagogy using COCA on different proficiency levels demonstrated that intermediate and advanced learners typically showed greater gains, though improvements were reported across all levels (Chen, 2017). Furthermore, some studies found that DDL using the COCA is effective for collocation learning in L2 writing, but higher proficiency learners performed better than lower proficiency learners. Lower-proficiency learners even perceived the corpus as less helpful than higher-proficiency learners (Wu, 2016). They noted improvements in L2 learners who integrated corpus tools into language learning, but the effectiveness varied based on learner proficiency, especially for lower-proficiency learners who often struggle with extracting meaningful patterns and applying them in productive tasks. The difficulty stems from the cognitive load involved in analyzing large datasets without sufficient guidance (Boulton, 2017).

Recent studies have shifted toward investigating how teacher instruction embedded by corpus tools specifically impacts verb-noun collocation learning for low-proficiency learners. The concept of CBLP emphasizes the integration of corpus technology into language learning through a strategically designed curriculum and instructional strategies, highlighting that beyond selecting an appropriate corpus tool effective pedagogy must include well-designed scaffolding mechanisms

such as guided discovery, explicit instruction, and corpus-based interactive activities (Ma et al., 2024). Rather than assuming that corpus exploration is inherently unsuitable for such learners, studies suggest that corpus tools must be integrated into teaching practices through scaffolded activities that progressively build learners' analytical skills (Alsehibany & Abdelhalim, 2023, p. 178). Accordingly, low-proficiency learners are not overwhelmed by the large corpus data but instead receive structured support that enhances their ability to extract and apply collocational patterns meaningfully. CBLP aligns with recent research findings that scaffolding plays a crucial role in mitigating the challenges of corpus-based learning for low-proficiency learners. In this regard, our study also explored how instructional strategies can maximize the benefits of corpus-based collocation learning, particularly with SkELL tool.

Notably, the absence of research on the use of SkELL for Chinese low-proficiency learners presents a significant gap in the existing literature. While SkELL has been studied in the context of Japanese learners (Hirata & Hirata, 2019), its applicability to Chinese students remains less explored. Moreover, the comparative effectiveness of SkELL and the COCA in supporting verb-noun collocation learning has yet to be systematically investigated. Therefore, this study seeks to address these gaps by investigating how SkELL-supported scaffolding interventions impact the enhancement of verb-noun collocations for low-proficiency Chinese learners.

Method

The current study aimed to compare the impact of SkELL and COCA on English verb-noun collocation learning for low-proficiency Chinese EFL learners. Therefore, an intervention design with pre- and post-tests was adopted to examine the differences among three groups, the SkELL group (n=20), COCA group (n=20), and control group (n=20).

Participants

This study was carried out at a vocational college during the fall semester of 2024 in southwestern China, where English is a compulsory course for undergraduate students. The 60 first-year EFL students, aged between 17 and 20, volunteered to participate in the study. They had studied English for six years but their English vocabulary size was estimated to be no more than 1000 words based on their average 50 out of 150 on the college entrance examination, which is equivalent to level 3 (or the lower stage) in China's Standards of English Language Ability (CSE) developed by the National Education Examinations Authority and Department of Language Information Management (Ministry of Education, 2018). A presurvey ensured that all the participants had no prior knowledge about corpus use in collocation learning.

To address our research questions on the differential impact of SkELL and COCA on English verb-noun collocation learning for low-proficiency Chinese EFL learners, we randomly divided the students into three groups (SkELL group, COCA group, and control group). The first group was the SkELL group, 20 participants who were instructed using the SkELL; the second was the COCA Group, 20 participants who were taught collocations through COCA; and the last was the control group (C-Group), 20 participants who received traditional lesson-based instruction.

The duration of the English course is 8 weeks with each lesson lasting for 60 minutes. The course aims to provide students with a basic understanding of corpora, and corpus tools, and develop students' overall English skills, focusing specially on verb-noun collocation learning. By the end of the course, SkELL group students were invited to evaluate the effectiveness of using

SkELL for verb-noun collocation learning through a survey. During the study, three groups were taught using the same teaching materials, textbook, and instructor.

A convenience sampling method was utilized to recruit participants for the questionnaire as well as the pre-and post-tests. Additionally, purposeful sampling was implemented to select interview participants based on their availability and relevance to the research objectives (Leedy & Ormrod, 2015).

Instruments

Two instruments were designed for data collection: pre- and post-collocation tests, and a survey. The instrument was checked by the three proficient EFL instructors and piloted to a group of preparatory class students including 20 students for validity and reliability.

Pre- and Post-tests

To assess the impact of different instructional methods on students' proficiency in verb-noun collocations, a pre-test was administered to all participants prior to the intervention, followed by a post-test at the end of the lesson. Both tests comprised 20 multiple-choice items and 20 gap-filling items, designed to evaluate students' collocational knowledge comprehensively. Each correct response was awarded one point, with a maximum achievable score of 40. The pre-test and post-test utilized the same set of multiple-choice and gap-filling items to ensure consistency and comparability of results (refer to Appendix A). These test formats were selected as they are among the most commonly used and effective measures for assessing collocational competence (Nguyen, 2022). A one-way analysis of variance (ANOVA) was run on the data obtained from the proficiency test.

Questionnaire

The basic assumption that underlies all attitude scales is that it is possible to discover attitudes by asking individuals to respond to a series of statements of preference (Fraenkel et al., 2012). To measure students' overall attitudes toward the use of SkELL for collocation learning, an online questionnaire with 16-item was developed by examination of previous literature with slight modification, e.g., Hirata and Hirata (2019), along with their considerations of the specific use of the corpus in the classes serving as research sites for the study (refer to Appendix B).

As the commonly used attitude scale in educational research is the Likert scale, this part consisted of 16 statements to which the respondents were asked to indicate their degree of agreement on a scale of 1-6 (1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree). For the analysis of this Likert-type data, the mean scores of the students' responses for each question were computed relative to the six-point scale employed. A standard deviation (± SD) was also obtained to examine statistically significant differences between students' opinions. In the present study, the data are presented as mean ± SD.

Procedures

To investigate the differential impact of SkELL and COCA on English verb-noun collocation learning for low-proficiency Chinese learners, an 8-lesson study was arranged as an intervention. To begin with, the pretest and the intervention tasks were piloted on 20 learners. The analysis of the data revealed that, as expected, learners did not know the targeted collocations and corpus use.

The design of the study is summarized in Table One.

Table 1. *The procedure of the study*

	Groups		
Lessons	SkELL group (n=20)	COCA group (n=20)	Control group (n=20)
1	Pretest (all groups) + briefing		
2-7	SkELL-based instruction	COCA-supported instruction	Lecture-based instruction
8	Post-test (all groups) + Questionnaire (corpus-based groups)		

In the first lesson, a pretest was administered in which students were asked to finish a pretest, including 20 multiple-choice items as a recognition measure and 20 gap-filling items as a productive measure within 30 minutes. After taking pretest, 6 physical lessons were conducted among SkELL group, COCA group and the control group respectively. In the last lesson, they were required to take the post-test, and complete survey question to explore their overall evaluation of using SkELL for collocation learning.

Target Collocations

During the 8-lesson intervention, students were taught 60 verb-noun collocations. These collocations were selected from the Academic Collocation List (ACL) created by Ackermann and Chen (2013), available at https://www.pearsonpte.com/teachers/academic-collocation (refer to Appendix C). The nodes and collocates were chosen in line with their textbook words in terms of the frequency level to measure their knowledge of the collocations.

Collocation Course Design

A 6-lesson collocation English course was held for each group at a Chinese vocational college. Each lesson took about 60 minutes in which 5-6 verb-noun collocations were presented to the learners under three different conditions. The control group was instructed on the target collocations through traditional lesson-based instruction. For example, the teacher first wrote each collocation on the board and then provided its meaning in Chinese followed by translating some typical example sentences. Then, the control group completed a fill-in-the-blank activity without any access to a dictionary. Their answers were checked by the teacher at the end. In contrast, the SkELL group and COCA group learners were asked to search for the given collocations in the concordance and read the example sentences. With the guidance of the teacher, they need to discuss how to identify the collocations in different contexts and summarize their general usage. The specific teaching activities for the collocation course design are shown in Appendix D.

This collocation course design embedded with SkELL and COCA was based on the four-step CBLP design model developed by Ma et al. (2022). In the first step, the lesson began with a brief explanation of verb-noun collocations in simple terms, ensuring that students understand the concept and raise their awareness of the target collocation features. To test their prior knowledge, the teacher provided relatable examples such as “make a fortune” instead of “get a fortune”. The importance of verb-noun collocations for natural and fluent English was emphasized, highlighting their role in effective communication. In the second step, students were guided to access the

SkELL or COCA website on their mobile phones. The teacher demonstrated each of their essential features, including the interfaces with word search function (see Figure 1), example sentences (see Figure 2), and concordance display (see Figure 3). Using a sample verb such as “achieve”, the teacher illustrated how to identify commonly associated nouns and how SkELL or COCA presented collocation patterns (see Figure 4 for COCA; see Figure 5 for SkELL). To reinforce learning, students work in pairs to practice using SkELL or COCA. They were given some common verb-noun collocations from the target list. Each pair followed structured guidelines with teacher guidance, ensuring students correctly interpreted the results from SkELL or COCA. In the third step, students were encouraged to summarize the rules that they explored based on their hands-on activities. The teacher concluded with brief feedback on the student’s exploration. In the last step, students individually completed fill-in-the-blank activities in the textbook. The teacher circulated among students, offering feedback and clarification where needed.

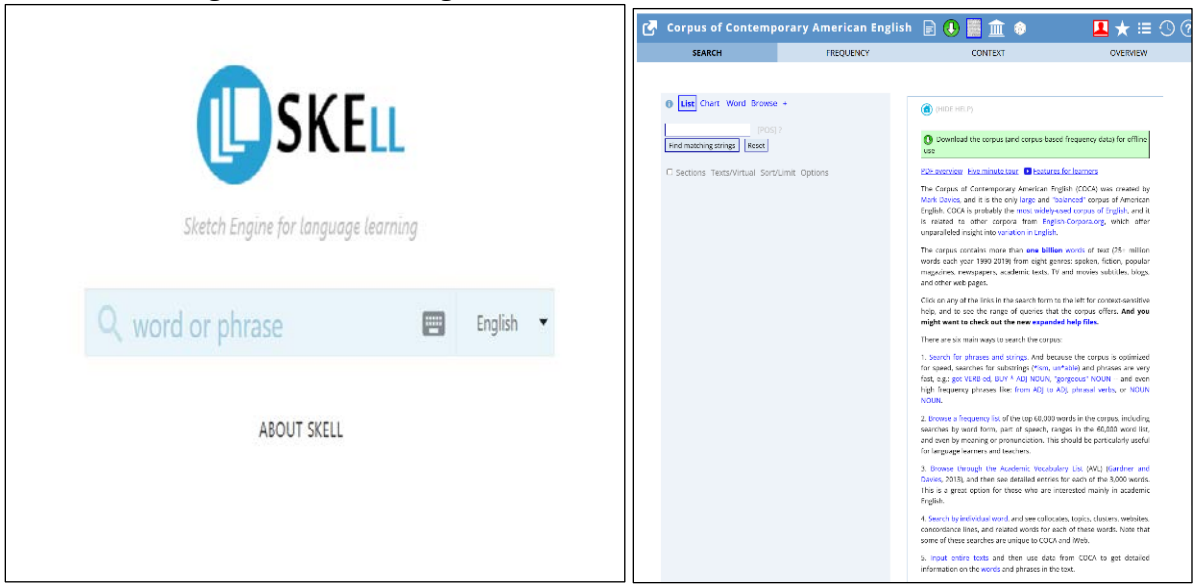


Figure 1. The interfaces of SkELL and COCA



Figure 2. SkELL example sentences display *achieve*

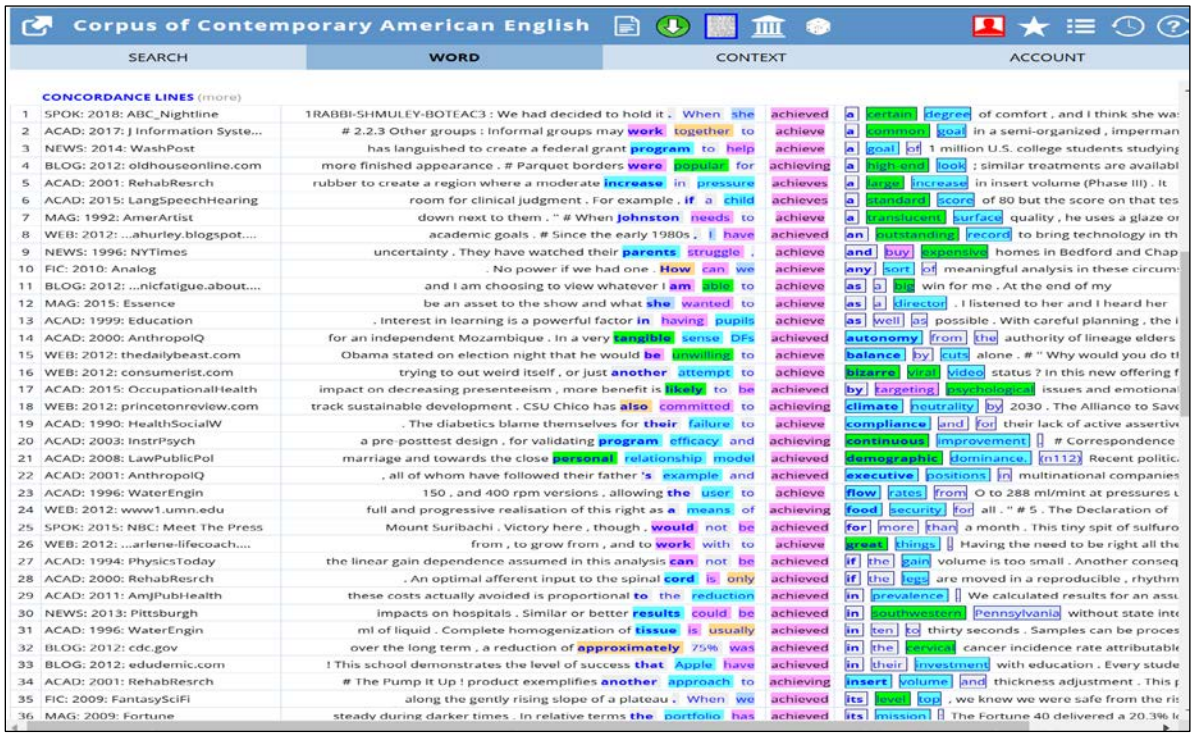


Figure 3. COCA Concordance Display of *achieve*

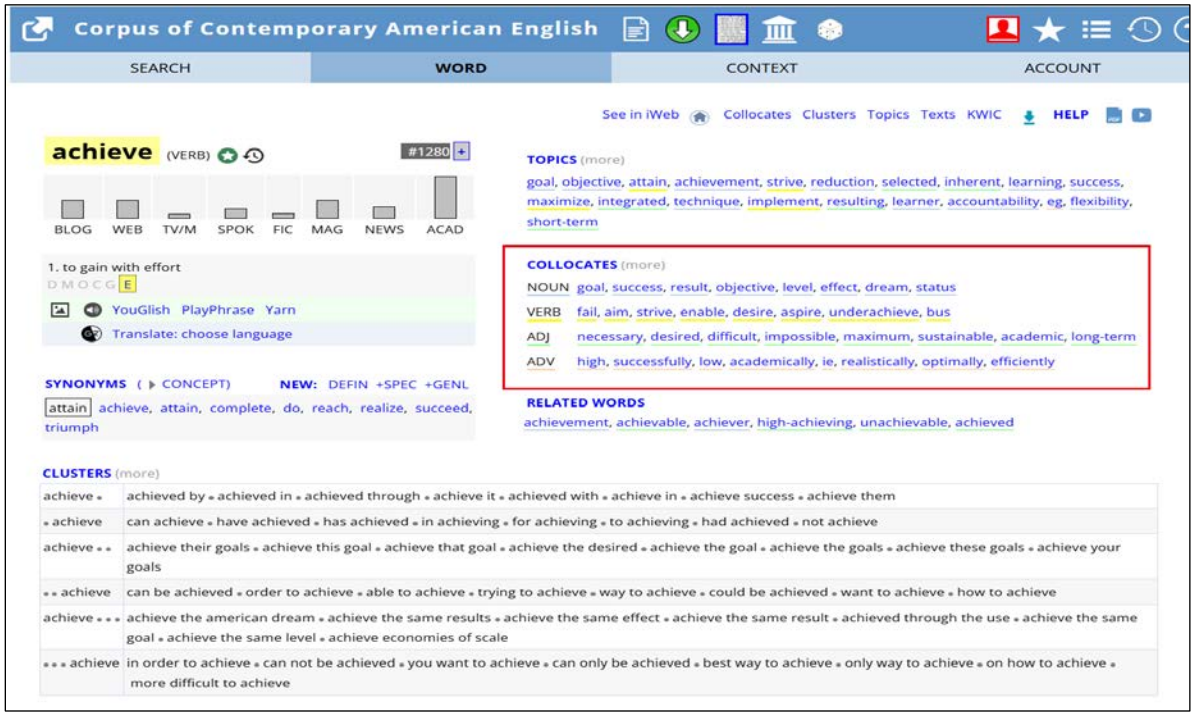


Figure 4. COCA Search Results for *achieve*

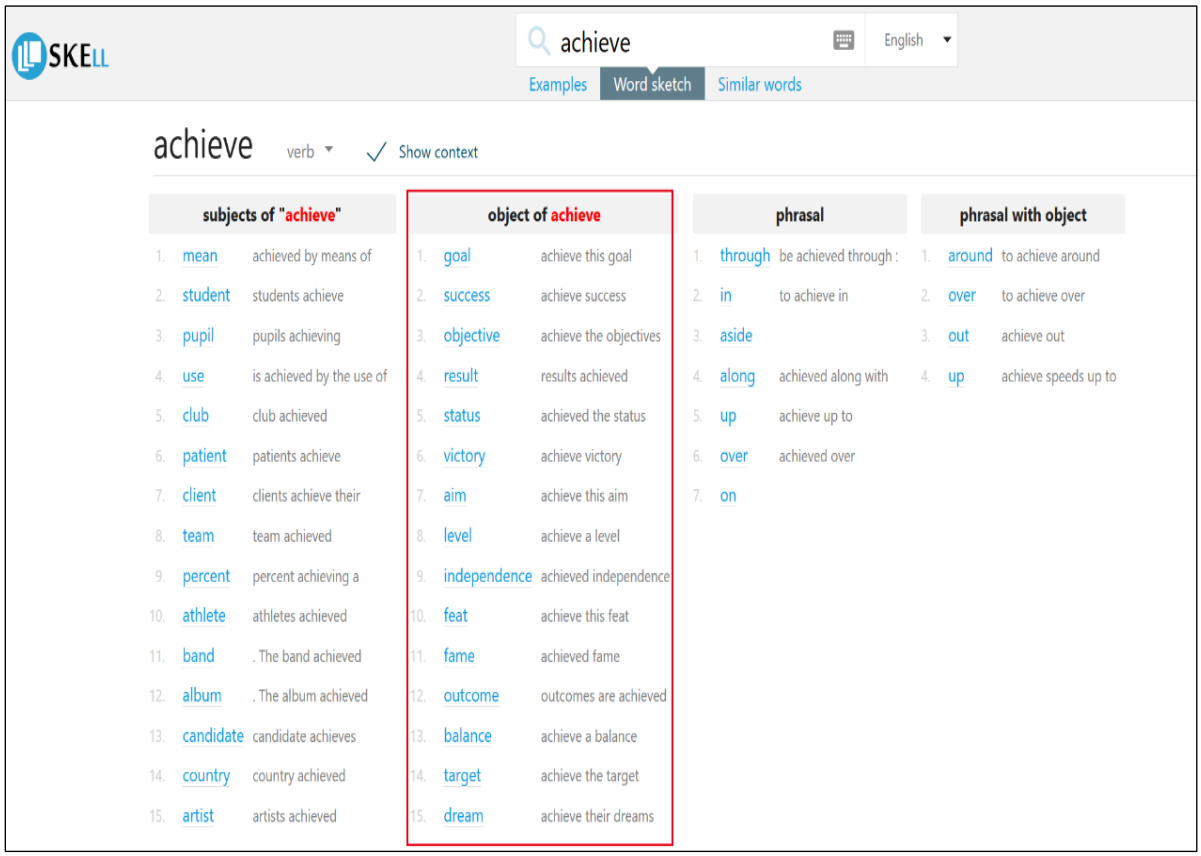


Figure 5. SkELL word sketch display of *achieve*

Scoring and Data Analysis

The pre- and post-tests had a maximum score of 40, with each correct response earning one point and incorrect responses receiving zero. The proficiency test results were analyzed using the Statistical Package for the Social Sciences (SPSS). A one-way analysis of variance (ANOVA) was conducted to confirm that the three participating groups were homogeneous before the intervention. Additionally, a repeated-measures ANOVA was performed to compare the group's performance on the pretest and post-test.

Results

Quantitative Data Analysis

Before conducting statistical comparisons of the pre-test and post-test scores among the three groups, it was essential to evaluate the normality of the data distribution. Assessing normality ensured that the assumptions underlying the parametric test were met, thereby validating the reliability of the statistical results. This study adopted the Shapiro-Wilk test which has been evaluated as the most powerful test for normality (Mbah & Paothong, 2015). The p-value is a key statistical measure that helps determine whether the data deviates significantly from normality. If $p > 0.05$, the data is normally distributed. In contrast, if $p \leq 0.05$, the data significantly deviates from normality (Mbah & Paothong, 2015).

Table Two presented the results of the Shapiro-Wilk tests of normality for the pretest and

post-test scores among COCA group (n=20), SkELL group (n=20) and control group (n=20), showing that all of the groups were normally distributed.

Table 2 *Shapiro-Wilk Tests of Normality*

		Shapiro-Wilk		
		Statistic	df	Sig.
Pretest	COCA	.970	20	.764
	SkELL	.940	20	.241
	Control	.934	20	.182
Post-test	COCA	.952	20	.393
	SkELL	.935	20	.189
	Control	.959	20	.526

Afterwards, Table Three showed the descriptive statistics for the pretest and post-test, revealing that the means of corpus-based groups (i.e., the SkELL group and the COCA group) were better than the means of no-corpus group (i.e., the control group). The results demonstrated the pretest scores across the three groups were relatively similar, suggesting that students had comparable collocation proficiency before the intervention. However, the results of the post-test indicated that students performed better after the intervention than before. It is worth noting that the SkELL group showed the highest improvement, with a mean increase of 11.6 points (ranging from 18.20 to 29.80 in mean scores).

Table 3. *Descriptives Statistics for Collocation Tests*

						95% Confidence Interval for Mean			
		N	Mean	SD	Std. Error	Lower Bound	Upper Bound	Min	Max
Pretest	COCA	20	16.85	5.412	1.210	14.32	19.38	6	26
	Control	20	17.85	5.669	1.268	15.20	20.50	6	25
	SkELL	20	18.20	4.895	1.094	15.91	20.49	7	25
	Total	60	17.63	5.275	.681	16.27	19.00	6	26
Posttest	COCA	20	25.65	7.989	1.786	21.91	29.39	12	38
	Control	20	22.45	7.824	1.749	18.79	26.11	8	35
	SkELL	20	29.80	5.709	1.277	27.13	32.47	20	38
	Total	60	25.97	7.744	1.000	23.97	27.97	8	38

Homogeneity of variances is an assumption for parametric tests. Table Four showed that all *p*-values were well above 0.05, meaning that the variances among the three groups in the pretest were homogeneous. Hence, a one-way ANOVA was run to compare group means.

Table 4. *Tests of Homogeneity of Variances*

		Levene Statistic	df1	df2	Sig.
Pretest	Based on Mean	.438	2	57	.647
	Based on Median	.316	2	57	.730
	Based on the Median and with adjusted df	.316	2	56.799	.730
	Based on trimmed mean	.417	2	57	.661
Post-test	Based on Mean	2.376	2	57	.102

Based on Median	1.863	2	57	.165
Based on Median and with adjusted df	1.863	2	53.686	.165
Based on trimmed mean	2.339	2	57	.106

In Table Five, the results of the one-way ANOVA for the pretest score demonstrated that there was no statistically significant difference among the three groups, F (2, 57) =0.345, p=0.71, η^2=0.012, while there was a significant difference for the post-test score among the three groups with a moderate effect size, F (2, 57) =5.17, p=0.009, η^2=0.154.

Table 5 *ANOVA Test of Pre- and Post-tests Score within Groups*

		Sum of Squares	df	Mean Square	F	Sig.	Eta-squared
Pretest	Between Groups	19.633	2	9.817	.345	.710	.012
	Within Groups	1622.300	57	28.461			
	Total	1641.933	59				
Post-test	Between Groups	543.233	2	271.617	5.170	.009	.154
	Within Groups	2994.700	57	52.539			
	Total	3537.933	59				

To locate the exact difference among groups, a Tukey HSD post-hoc analyse was run. Results of the post-test score showed that there was a significant difference between the SkELL group and the control group in learning verb-noun collocations, *p*=0.006. However, the main effect between the COCA group and either the control (*p*=0.35) or SkELL (*p*=0.1175) groups did not reach statistical significance, suggesting that the COCA group's learning method did not lead to a statistically significant improvement. The results of the Tukey HSD post-hoc test further demonstrated that the SkELL group's learning method was effective in improving students' verb-noun collocation knowledge, as shown in Table Six.

Based on these results, the answers to the research question 1 and 2 are positive. We can conclude that SkELL was the most effective corpus tool for learning verb-noun collocations among low-proficiency students, even though corpus-based instruction performed better than traditional lesson-based instruction.

Table 6. *Tukey HSD Post Hoc Test for Multiple Comparisons*

Dependent Variable			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Post-test	COCA	Control	3.200	2.292	.350	-2.32	8.72
		SkELL	-4.150	2.292	.175	-9.67	1.37
	Control	COCA	-3.200	2.292	.350	-8.72	2.32
		SkELL	-7.350*	2.292	.006	-12.87	-1.83
	SkELL	COCA	4.150	2.292	.175	-1.37	9.67
		Control	7.350*	2.292	.006	1.83	12.87

*. The mean difference is significant at the 0.05 level.

Survey Analysis

The questionnaire was designed and administered to 20 participants after the intervention to evaluate students' perceptions of SkELL-supported verb-noun collocation learning. It contained 16 items, each measuring different aspects of students' experiences and attitudes. To ensure the

questionnaire's internal consistency, Cronbach's Alpha was calculated, yielding a value of 0.765. This indicated acceptable reliability, suggesting that the items collectively measure a consistent construct (see Table Seven).

The overall mean score across all items was 5.20 out of 6, reflecting a positive perception of SkELL-supported collocation learning. The highest-rated item was Item 5 (Mean = 5.70, SD = 0.657), indicating that most students had never used SkELL or online dictionaries before the course. This finding was significant because it suggested that students' positive feedback was not biased by prior familiarity, but rather by their new learning experience with SkELL. Interestingly, students strongly expressed that SkELL should be widely introduced in colleges after use, according to Item 14 (Mean = 5.25, SD = 0.550), suggesting that students found SkELL useful, accessible, and beneficial for learning collocations, reinforcing its practicality as an educational tool. As shown in Item 11, SkELL provided appropriate example sentence length, making it easier for students to read. While most items were rated positively, some had lower mean scores and higher SD values, indicating more diverse opinions. Item 16 (Mean = 4.85, SD = 0.875) and Item 15 (Mean = 4.55, SD = 0.887) received the lowest ratings, implying that students were less confident about SkELL's role in improving overall other aspects of language learning, like writing and reading skills. The higher SD values suggested mixed perceptions, where some students may have found SkELL effective for collocations but not necessarily for broader application in other aspects of language learning. Therefore, the results highlighted a potential demand—while SkELL helped with collocation knowledge, students might still require additional writing support or guidance to fully integrate corpus use into their writing in future learning.

Table 7. *Students' Assessment of SkELL-supported Collocation Learning*

Item No.	Items	Mean	SD	Cronbach's alpha
1	SkELL's large volume of example sentences is more useful for your English study than those presented by English dictionaries.	5.25	.550	
2	SkELL assists in learning more easily how words and phrase's function than English dictionaries presenting a limited number of example sentences.	5.30	.571	
3	SkELL is better than English dictionaries because it enables you to search words and/or phrases.	5.25	.550	
4	Using both SkELL and English dictionaries is indispensable for your English study.	5.30	.571	
5	I hadn't used either SkELL or online English dictionaries until I took this course.	5.70	.657	
6	SkELL helps you learn various collocations.	5.25	.550	
7	SkELL is useful when you prepare for English exams.	5.35	.489	
8	SkELL enhances your motivation to study English.	5.20	.523	
9	The number of example sentences provided by SkELL is appropriate for my English study.	5.10	.641	
10	SkELL attracts and interests me, so studying with it is fun and entertaining.	5.25	.444	
11	I don't need longer example sentences in SkELL.	5.20	.616	
12	The layout of the website and interface is easy to access via mobile phones.	5.55	.510	

13	Using SkELL in the classroom helped me study English successfully.	4.85	.489	
14	SkELL should be widely used in college school.	5.25	.550	
15	SkELL is effective in improving reading skills.	4.55	.887	
16	SkELL is effective in improving writing skills.	4.85	.875	
	Total	5.20	.604	0.765

Discussion

Based on the results reported above, three groups presented different effects on learning verb-noun collocations. First, the non-significant pre-test results ($F = 0.345$, $p = 0.710$) indicated that all participants started at a similar level of collocation proficiency, ensuring the validity of the experimental design. In contrast, the post-test results ($F = 5.170$, $p = 0.009$) demonstrated a statistically significant difference among the groups, indicating that at least one instructional method led to substantial improvements in collocation proficiency. Consequently, corpus-based learning may enhance students' ability to acquire and use verb-noun collocations more effectively compared to traditional lesson-based instruction. Second, to further explore the exact differences among groups, a Tukey HSD post-hoc analysis was used to identify which groups benefited the most. Results claimed that SkELL-supported instruction outperformed both the COCA-supported and control groups, revealing the significant impact of SkELL-supported learning on verb-noun collocation proficiency for low-proficiency students. All in all, the findings suggested that learners benefited most from SkELL-supported learning when provided with well-designed scaffolding.

In addition to the pre-and post-tests results, this study also explored learners' perceptions and assessment of SkELL-supported collocation learning after the intervention, utilizing a 16-item questionnaire to collect participants' feedback. The overall results indicated that learners held a positive attitude toward SkELL-supported collocation learning, suggesting that students found SkELL engaging, useful, and beneficial in enhancing their collocation knowledge.

It was worth noting that these findings contrasted with previous studies (Boulton & Cobb, 2017; Daskalovska, 2015; Huang, 2014), which suggested that corpus-based learning may not be suitable for lower-level learners due to the overwhelming amount of information available in corpora, easily leading them to confusion and cognitive overload. However, they aligned with some recent studies (Liu & Ma, 2025; Ma et al., 2024; Saeedakhtar et al., 2020), which highlighted the effectiveness of corpus-based learning for low-proficiency learners. Most importantly, a key contribution of this study was its demonstration that the effectiveness of corpus-based learning for low-proficiency students largely depended on the selection and application of an appropriate corpus tool with well-designed scaffoldings. While corpus-based learning, in general, proved more beneficial than traditional instruction, this study revealed that SkELL-supported verb-noun collocation learning outperformed COCA-supported verb-noun collocation learning, which suggested that SkELL was a more effective and user-friendly tool for low-proficiency students. Unlike more complex corpus tools, for example, COCA, which may present too much unfiltered linguistic data, SkELL provided simplified, high-frequency collocation patterns that tailored to those who possessed limited linguistic proficiency. Hence, this study emphasized that not all corpus tools were equally beneficial for all learners. The superior performance of the SkELL group compared to the COCA group suggested that corpus tools specifically tailored to students' proficiency levels, when integrated with pedagogical scaffolding, are more effective for low-

proficiency learners.

Consequently, while the results of this study reinforced the significance of level-matching corpus tool for low-proficiency learners, they also underscored the importance of pedagogical scaffolding for corpus-based learning. Nevertheless, it is important to note that while students appreciated SkELL for collocation learning, they were less confident about its impact on their other aspects of language learning (i.e., reading and writing). Thus, it highlighted the need for teacher-mediated corpus use, where teachers provide effective scaffolding, ensuring that students can make full use of the corpus data.

This study provided valuable insights, yet it has some limitations. The sample size was relatively small (N = 20), which may restrict the generalizability of the findings. Additionally, the study focused on only one type of collocation, limiting the scope of its applicability to broader linguistic contexts. Future research should consider longitudinal studies for a larger sample to explore the effectiveness of SkELL-based instruction. Furthermore, it can also investigate how customized corpus-based instruction can enhance students' ability to improve overall language proficiency.

Conclusion

While early studies (Boulton & Cobb, 2017; Yoon, 2008) suggest that corpus tools may not be well-suited for low-proficiency learners, recent studies (Chen, 2017; Gablasova et al., 2017; Hirata & Hirata, 2019; Liu & Ma, 2025; Ma et al., 2024) challenge this notion by emphasizing the role of scaffolding and user-friendly tools. The concept of CBLP further reinforces that effective corpus-based learning is not solely about tool selection but also involves strategic curriculum design and instructional interventions (Ma et al., 2024; Ma et al., 2022). Our study builds upon these findings by demonstrating that with appropriate instructional support, corpus-based learning can be highly effective for low-proficiency learners, particularly when using simplified tools like SkELL. The study contributes to the growing body of literature advocating for a more nuanced approach to corpus-based learning. While past research has established the general effectiveness of corpus tools in L2 learning, our study highlights the importance of tool selection and instructional design in optimizing collocation acquisition for low-proficiency learners.

Overall, the findings examined the impact of the corpus tool on verb-noun collocation learning for Chinese low-proficiency students and further explored the exact difference between SkELL- and COCA-supported verb-noun collocations learning under the same teacher scaffolding. The findings highlighted the potential force of SkELL-supported verb-noun collocation learning in enhancing verb-noun collocation proficiency for low-proficiency students. The results of the questionnaire showed students held a positive attitude toward the SkELL corpus, due to its manageable example sentence length that was easy to read. Therefore, this study carries important pedagogical implications for the implementation of SkELL and exploring other corpus-based tools for improving language proficiency. Hence, it is recommended that the SkELL corpus tool be introduced into Chinese EFL low-proficiency learners' classes, especially in vocational college settings, to facilitate verb-noun collocational learning.

By addressing the lack of research on SkELL for Chinese learners, we contribute to expanding the scope of corpus-based collocation instruction in diverse linguistic contexts. Future research should explore additional scaffolding mechanisms and investigate the long-term retention effects of corpus-based collocation learning. Additionally, further comparative studies between

different corpus tools can provide deeper insights into the affordances and limitations of various platforms.

Funding: This research is not funded.

Acknowledgement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: During the preparation of this work, the authors used Grammarly to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and they take full responsibility for the content of the publication.

About the authors:

Shuang Yang is currently a PhD candidate in TESL at the Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia. Her main research interests include applications of corpus linguistics in language learning and teaching, L2 teaching pedagogy, and computer-assisted language learning (CALL). ORCID iD: https://orcid.org/0009-0006-7581-1571. Email: yangshuang@graduate.utm.my

Wan Nur Asyura Wan Adnan is a Senior Lecturer at the Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia. Her main research interest is 21st-century language teaching methodology and discourse analysis. ORCID ID: https://orcid.org/0000-0002-2332-6703. Email: asyura.kl@utm.my

References

Ackermann, K., & Chen, Y. (2013). Developing the Academic Collocation List (ACL)–A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes, 12*(4), 235-247. https://doi.org/10.1016/j.system.2017.08.004

Akbari, J., Haghverdi, H., & Biria, R. (2015). Instructional efficacy of corpus-based tools in teaching collocations to Iranian university students with different majors. *Journal of Applied Linguistics and Language Research, 2*(8), 218-229.

Akıncı, M., & Yıldız, S. (2017). Effectiveness of corpus consultation in teaching verb+ noun collocations to advanced ELT students. *Eurasian Journal of Applied Linguistics, 3*(1), 91-108. https://doi.org/10.32601/ejal.461036

Al-Mahbashi, A., Noor, N. M., & Amir, Z. (2015). The effect of data driven learning on receptive vocabulary knowledge of Yemeni University learners. *3L, Language, Linguistics, Literature, 21*(3), 13-24.

Alharbi, M. S. (2024). Unveiling linguistic frontiers in the 21 st century: A systematic literature review of the rise of big data's role in data-driven language learning in the Saudi Arabian context. *Arab World English Journal, 15*(4), 3-21.

https://doi.org/10.24093/awej/vol15no4.1
Alsehibany, R. A., & Abdelhalim, S. M. (2023). Enhancing academic writing vocabulary use through direct corpus consultation: Saudi english majors' perceptions and experiences. *Arab World English Journal*, *14*(1), 178. https://doi.org/10.2139/ssrn.4431295
Altun, H. (2021). The learning effect of corpora on strong and weak collocations: implications for corpus-based assessment of collocation competence. *International Journal of Assessment Tools in Education*, *8*(3), 509-526. https://doi.org/10.21449/ijate.845051
Ayyat, A. (2024). Implementing a three-stage approach to teach writing skills using the corpus of contemporary American English in a foundation program context. *Arab World English Journal*, *15*(4), 241-254. https://doi.org/10.24093/awej/vol15no4.15
Boers, F., Lindstromberg, S., & Eyckmans, J. (2014). Some explanations for the slow acquisition of L2 collocations. *Vigo International Journal of Applied Linguistics*(11), 41-62.
Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348-393. https://doi.org/10.1111/lang.12224
Chen, L. (2017). Corpus-aided Business English collocation pedagogy: An empirical study in Chinese EFL learners. *English Language Teaching*, *10*(9), 181-197. https://doi.org/doi.org/10.5539/elt.v10n9p181
Crosthwaite, P. (2024). *Corpora for language learning*. Routledge.
Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, *28*(2), 130-144. https://doi.org/10.1080/09588221.2013.803982
Durmuş, B. Y., & Kilimci, A. (2020). A corpus-driven deductive and inductive collocation instruction for young learners of English. *International Journal of Language Academy*, *8*(4), 371-383. https://doi.org/10.29228/ijla.46179
Fraenkel, J., Wallen, N., & Hyun, H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill Education.
Frankenberg-Garcia, A. (2024). DDL support for academic English collocations. In *Corpora for language learning* (pp. 71-83). Routledge. https://doi.org/10.4324/9781003413301-6
Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, *67*(S1), 155-179. https://doi.org/10.1111/lang.12225
Girgin, U. (2019). The effectiveness of using corpus-based activities on the learning of some phrasal-prepositional Verbs. *Turkish Online Journal of Educational Technology-TOJET*, *18*(1), 118-125.
Hirata, Y., & Hirata, Y. (2019). Applying 'Sketch Engine for language learning' in the Japanese English classroom. *Journal of Computing in Higher Education*, *31*(2), 233-248. https://doi.org/10.1007/s12528-019-09208-z
Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, *26*(2), 163-183. https://doi.org/10.1017/s0958344014000020
Jafarpour, A. A., Hashemian, M., & Alipour, S. (2013). A corpus-based approach toward teaching collocation of synonyms. *Theory & Practice in Language Studies (TPLS)*, *3*(1), 51-60. https://doi.org/10.4304/tpls.3.1.51-60

Kilgarriff, A., Marcowitz, F., Smith, S., & Thomas, J. (2015). Corpora and language learning with the Sketch Engine and SKELL. *Revue Française de Linguistique Appliquée, 20*(1), 61-80. https://doi.org/10.3917/rfla.201.0061

Leedy, P. D., & Ormrod, J. E. (2015). *Practical research*. Pearson Education Limited.

Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology, 21*(3), 153-171.

Liu, J., & Ma, Q. (2025). Examining corpus-based language pedagogy (CBLP) practices in data-driven learning (DDL) for low-proficiency L2 English learners: A meta-analysis. *Educational Technology & Society, 28*(2), 53-76. https://doi.org/10.30191/ETS.202504_28(2).RP04

Liu, T., & Gablasova, D. (2023). Data-driven learning of collocations by Chinese learners of English: a longitudinal perspective. *Computer Assisted Language Learning*, 1-26. https://doi.org/10.1080/09588221.2023.2214605

Łuszcz, M. (2016). Enhancing development of communicative skills in English through teaching English collocations in a corpus-based approach. *Kalbų Studijos*(29), 88-101.

Ma, Q., & Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *Journal of China Computer-Assisted Language Learning, 1*(1), 177-190. https://doi.org/10.1515/jccall-2021-2008

Ma, Q., Mei, F., & Qian, B. (2024). Exploring EFL students' pronunciation learning supported by corpus-based language pedagogy. *Computer Assisted Language Learning*, 1-27. https://doi.org/10.1080/09588221.2024.2432965

Ma, Q., Tang, J., & Lin, S. (2022). The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning, 35*(9), 2731-2760. https://doi.org/10.1080/09588221.2021.1895225

Mbah, A. K., & Paothong, A. (2015). Shapiro–Francia test compared to other normality test using expected p-value. *Journal of Statistical Computation and Simulation, 85*(15), 3002-3016.

Men, H. (2020). Data-driven learning in enhancing learners' language idiomaticity. *International Journal of Emerging Technologies in Learning (iJET), 15*(23), 27-41. https://doi.org/10.3991/ijet.v15i23.19023

Ministry of Education (2018). *The China's standard of English ability was officially released by the ministry of education and the national language committee*. http://www. moe. gov.cn/srcsite/A19/s229/201804/t20180416_333315.html

Nguyen, H. (2022). *New measures of academic collocation knowledge: Wordlist-based test development and argument-based validation* [Thesis, Victoria University of Wellington Library.

Saeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of hands-on and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. *System, 91*, 102268. https://doi.org/10.1016/j.system.2020.102268

Satake, Y. (2023). The effects of corpus use on learning L2 collocations of light verbs and nouns. *Asia Pacific Journal of Corpus Research, 4*(2), 41-55.

Sun, W., & Park, E. (2023). EFL learners' collocation acquisition and learning in corpus-based

instruction: A systematic review. *Sustainability*, *15*, 13242. https://doi.org/10.3390/su151713242

Tabib, F. M. (2022). Exploring the effect of instructional scaffolding on foundation level students' writing at the city University College of Ajman: A case study. *Arab World English Journal, 13*(3), 185-201. https://doi.org/10.24093/awej/vol13no3.12

Ucar, S., & Yükselir, C. (2015). The effect of corpus-based activities on verb-noun collocations in EFL classes. *Turkish Online Journal of Educational Technology-TOJET, 14*(2), 195-205.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard University Press.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. https://doi.org/10.1017/cbo9780511519772

Wu, Y. A. (2015). Utilizing corpus resources accompanied by other consultation resources in enhancing collocation accuracy and collocation richness in L2 writing. [Doctoral thesis, University of California, Santa Barbara]. https://alexandria.ucsb.edu/lib/ark:/48907/f3kh0kh0

Wu, Y. A. (2016). The effects of utilizing corpus resources to correct collocation errors in L2 writing – Students' performance, corpus use and perceptions. In S. Papadima-Sophocleous, Bradley, L., & Thouësny, S. (Eds.), *CALL Communities and Culture – Short Papers from EUROCALL 2016* (pp. 479-484). Research-publishing.net. https://doi.org/10.14705/rpnet.2016.eurocall2016.610

Wu, Y. A. (2021). Discovering collocations via data-driven learning in L2 writing. *Language Learning & Technology, 25*(2), 192–214. https://doi.org/10125/73440

Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning & Technology, 12*(2), 31-48.

Appendices

Appendix A: Verb-noun Collocation Pre- and Post-tests

Verb-noun Collocation Pretest

Time allowed: 60 minutes
Total points: 40 (1 point per question)

Part A: Multiple Choice Questions (20 points)
Select the most appropriate verb that collocates with the underlined noun.
1. Students need to _____ practical experience before entering the workforce.
a) make
b) gain
c) take
d) do

2. The committee will _____ responsibility for organizing the conference.
a) assume
b) take
c) hold

d) both a and b

3. It is important to _____ precautions when conducting laboratory experiments.
a) make
b) take
c) do
d) have

4. The professor helped students _____ awareness of cultural differences.
a) raise
b) increase
c) develop
d) both a and c

5. The university will _____ measures to improve student satisfaction.
a) take
b) implement
c) adopt
d) all of the above

6. Participants must _____ permission before sharing research data.
a) get
b) obtain
c) receive
d) all of the above

7. The study aims to _____ progress in understanding climate change.
a) achieve
b) make
c) show
d) both a and b

8. Students should _____ notes during lectures.
a) write
b) take
c) make
d) both b and c

9. The research team will _____ recommendations based on their findings.
a) make
b) provide
c) give
d) both a and b

10. Universities must _____ standards of academic integrity.
a) maintain
b) uphold
c) keep
d) all of the above

11. Students need to _____ concentration during exams.
a) maintain

b) keep
c) hold
d) both a and b

12. The department will _____ changes to the curriculum next year.
a) make
b) implement
c) introduce
d) all of the above

13. Researchers should _____ care when handling sensitive data.
a) take
b) exercise
c) show
d) both a and b

14. The study will _____ light on this complex issue.
a) shed
b) throw
c) cast
d) all of the above

15. Students must _____ deadlines for assignment submission.
a) meet
b) follow
c) observe
d) both a and c

16. The university aims to _____ opportunities for international exchange.
a) create
b) provide
c) offer
d) all of the above

17. Professors should _____ feedback on student assignments.
a) give
b) provide
c) offer
d) all of the above

18. The research will _____ focus on qualitative methods.
a) place
b) put
c) have
d) both a and b

19. Students need to _____ respect for diverse viewpoints.
a) show
b) demonstrate
c) express
d) both a and b

20. The department will _____ priority to improving research facilities.
a) give
b) assign
c) set
d) all of the above

Part B: Gap-Filling (20 points)
Complete each sentence with an appropriate verb using the given initial letters.
1. Students must co_____ research before writing their papers.
2. The university will im_____ new policies next semester.
3. Researchers need to ga_____ data through surveys.
4. Students should ta_____ advantage of learning resources.
5. The department will ma_____ improvements to facilities.
6. Professors must pr_____ feedback on assignments.
7. Students need to de_____ skills in academic writing.
8. The university will of_____ support to international students.
9. Researchers should dr_____ conclusions from evidence.
10. Students must me_____ requirements for graduation.
11. The committee will se_____ priorities for the year.
12. Professors should ra_____ awareness about plagiarism.
13. Students need to ke_____ records of their research.
14. The department will es_____ guidelines for research.
15. Researchers must co_____ analysis of results.
16. Students should ex_____ caution when citing sources.
17. The university will pr_____ opportunities for learning.
18. Professors must ma_____ standards in assessment.
19. Students need to sh_____ progress in their studies.
20. Researchers should ob_____ protocols for data collection

Academic Verb-noun Collocation Post-test

Time allowed: 60 minutes
Total points: 40 (1 point per question)

Part A: Multiple Choice (20 points)
Choose the most appropriate verb that collocates with the underlined noun in academic contexts.
1. Researchers must _____ reliable data through systematic observation.
a) collect
b) gather
c) obtain
d) all of the above

2. Their methodology failed to _____ conclusive evidence for the hypothesis.
a) create
b) yield
c) generate
d) both b and c

3. The study aims to _____ a comprehensive analysis of socioeconomic factors.

a) do
b) make
c) conduct
d) perform

4. This theoretical framework helps scholars _____ a deeper understanding of the phenomenon.
a) achieve
b) gain
c) obtain
d) all of the above

5. The researcher needs to _____ significant differences between control groups.
a) identify
b) find
c) detect
d) both a and c

6. These findings _____ important implications for clinical practice.
a) carry
b) have
c) hold
d) all of the above

7. Their experiment failed to _____ support for the alternative hypothesis.
a) provide
b) give
c) lend
d) both a and c

8. The study will _____ particular emphasis on methodological innovations.
a) put
b) place
c) make
d) both a and b

9. Scholars must _____ careful consideration to potential confounding variables.
a) pay
b) give
c) take
d) both a and b

10. The results _____ light on previously unexplored aspects of cognitive development.
a) throw
b) shed
c) cast
d) all of the above

11. Researchers must _____ access to restricted demographic data.
a) get
b) gain
c) obtain
d) all of the above

12. The findings _____ significance beyond the immediate research context.
a) carry
b) bear
c) hold
d) both a and b

13. Scientists need to _____ caution when interpreting correlational data.
a) take
b) exercise
c) practice
d) both a and b

14. The analysis aims to _____ relationships between variables.
a) establish
b) determine
c) identify
d) all of the above

15. This framework helps _____ distinctions between similar concepts.
a) draw
b) make
c) establish
d) both a and c

16. The study will _____ reference to established theoretical models.
a) make
b) provide
c) give
d) both a and c

17. Researchers must _____ attention to methodological limitations.
a) pay
b) give
c) direct
d) all of the above

18. Their findings _____ substantial contribution to existing knowledge.
a) make
b) provide
c) offer
d) all of the above

19. The analysis will _____ precedence to quantitative measures.
a) give
b) take
c) grant
d) both a and c

20. These results _____ influence on future research directions.
a) have
b) exert

c) exercise
d) both a and b

Part B: Gap-Filling (20 points)
Complete each sentence with an appropriate verb that collocates with the given noun. Use only one word for each gap.
1. Students must co_____ extensive research before writing their dissertation.
2. Researchers need to de_____ clear evidence to support their hypotheses.
3. The professor will pr_____ detailed feedback on your assignments.
4. Scientists must es_____ reliable methods for data collection.
5. Students should ma_____ regular attendance throughout the semester.
6. Researchers must im_____ appropriate controls in their experiments.
7. The university will in_____ new policies next semester.
8. Scholars need to ge_____ original ideas for their research proposals.
9. Students must su_____ assignments by the stated deadline.
10. Researchers should ob_____ strict protocols during experiments.
11. The committee will re_____ applications for research funding.
12. Students need to de_____ critical thinking skills.
13. Professors must ad_____ effective teaching strategies.
14. Researchers should fo_____ clear research questions.
15. Students must ac_____ necessary skills for their field.
16. The department will or_____ academic conferences annually.
17. Researchers need to an_____ data systematically.
18. Students should ex_____ caution when citing sources.
19. Professors must en_____ academic integrity in assessments.
20. The framework helps ac_____ understanding of social processes.

Appendix B: Students' Questionnaire

Please answer all questions in the questionnaire.
The following questions are regarding your opinions on using SKELL corpus. Please use the scale below to circle the response that most closely resembles your perspectives (6-point Likert-scale questions: 1= strongly disagree; 2= disagree; 3= somewhat disagree; 4= somewhat agree; 5= agree; 6: strongly agree).

1. SkELL's large volume of example sentences is more useful for your English study than those presented by English dictionaries.	1	2	3	4	5	6
2. SkELL assists in learning more easily how words and phrase's function than English dictionaries presenting a limited number of example sentences.	1	2	3	4	5	6
3. SkELL is better than English dictionaries because it enables you to search words and/or phrases.	1	2	3	4	5	6
4. Using both SkELL and English dictionaries is indispensable for your English study.	1	2	3	4	5	6
5. I hadn't used either SkELL or online English dictionaries until I took this course.	1	2	3	4	5	6
6. SkELL helps you learn various collocations.	1	2	3	4	5	6
7. SkELL is useful when you prepare for English exams.	1	2	3	4	5	6

8. SkELL enhances your motivation to study English.	1	2	3	4	5	6
9. The number of example sentences provided by SkELL is appropriate for my English study.	1	2	3	4	5	6
10. SkELL attracts and interests me, so studying with it is fun and entertaining.	1	2	3	4	5	6
11. I need longer example sentences in SkELL.	1	2	3	4	5	6
12. The layout of the website and interface is easy to access via mobile phones.	1	2	3	4	5	6
13. Using SkELL in the classroom helped me study English successfully.	1	2	3	4	5	6
14. SkELL should be widely used in college school.	1	2	3	4	5	6
15. SkELL is effective in improving reading skills.	1	2	3	4	5	6
16. SkELL is effective in improving writing skills.	1	2	3	4	5	6

Appendix C: Target Collocation List

#	Component I	POS I	Component II	POS II	Addition
1	achieve (a)	v	goal	n	
2	achieve (an)	v	outcome	n	
3	acquire	v	knowledge	n	
4	address (an)	v	issue	n	
5	apply (a)	v	method	n	
6	bear	v	resemblance	n	(to)
7	carry out	v	research	n	(of)
8	cause	v	consequences	n	
9	collect	v	data	n	
10	convey (a)	v	message	n	
11	create	v	opportunities	n	
12	draw	v	attention	n	(to)
13	enhance	v	learning	n	
14	experience	v	difficulties	n	
15	face (a)	v	dilemma	n	
16	follow (a)	v	procedure	n	
17	gain	v	access	n	(to)
18	gain	v	insight	n	(into)
19	give	v	access	n	(to)
20	give	v	emphasis	n	
21	give	v	priority	n	(to)
22	give (a)	v	presentation	n	
23	give (an)	v	explanation	n	
24	have	v	access	n	(to)
25	have (a)	v	tendency	n	(to)
26	identify	v	factors	n	
27	identify (an)	v	issue	n	
28	increase	v	awareness	n	
29	lead to (the)	v	conclusion	n	

30	make	v	adjustments	n	
31	make (a)	v	contribution	n	
32	make (a)	v	living	n	
33	make (an)	v	assessment	n	
34	meet	v	criteria	n	
35	meet	v	expectations	n	
36	obtain (a)	v	result	n	
37	offer (an)	v	opportunity	n	
38	play (a)	v	role	n	(in)
39	pose (a)	v	problem	n	
40	pose (a)	v	threat	n	(to)
41	present	v	evidence	n	
42	provide	v	assistance	n	
43	provide	v	guidance	n	
44	provide	v	resources	n	
45	provide (a)	v	service	n	
46	provide (an)	v	opportunity	n	
47	raise	v	awareness	n	
48	reach (a)	v	consensus	n	
49	reach (a)	v	peak	n	
50	reach (an)	v	agreement	n	
51	receive	v	treatment	n	
52	report	v	findings	n	
53	require	v	knowledge	n	
54	seek	v	help	n	
55	set (a)	v	target	n	
56	show (a)	v	trend	n	
57	store	v	data	n	
58	take	v	responsibility	n	
59	take into	v	consideration	n	
60	undertake	v	work	n	

Appendix D: Collocation Course Design based on SkELL and COCA

Step	Objectives	Activities based on SkELL & COCA

Step 1: Test Prior Knowledge	Assessing students' prior knowledge and raise their awareness of the target collocation features	Ask students to fill in the blanks by using "do" or "make". • You have to make a living and if you're smart you can _____ a fortune in a short time. a. make b. take c. get d. do • If people stare at the sun long enough, they will _____blind because our eyes are very sensitive. a. get b. have c. take d. go • If the boss wanted to see him, he was probably going to _____ fired. a. go b. get c. come d. take
Step 2: Hands-on Search	Introducing SkELL and COCA; guiding students to access corpus tools; demonstrating basic SkELL and COCA features	Ask students to fill in the blank using SkELL and COCA • working together to _____ a common goal is important. • Every American president in the last 30 years has tried to _____a deal with Iran. • Heavy smokers lose their health. The are more likely to _____a disease of lung cancer in the future. Students need to based on the following guidelines: • Search for each pairs in SkELL or COCA. • Identify and list three verbs frequently used with the noun. • Construct one example sentence using each collocation.
Step 3: Inductive Discovery	Encouraging students to infer rules or patterns from corpus data	Ask students to summarize the language patterns. Usage of the verb-noun collocations: "achieve" often comes before noun like "goal" or "success" or "objective", but "goal" was the most frequent object of "achieve".
Step 4: Output Exercises	Applying learned target verb-noun collocation knowledge through tasks	Ask students to complete fill-in-the-blank activities with the help of corpus tools and explain the reasons for answers. The students were guided by the following points: • Use SkELL to identify suitable verb collocates for each noun or noun collocates after verb. • Fill in blanks in given sentences with the correct collocations. • Create their own example sentences using the identified verb-noun combinations.