

Corpus-based Instruction of Gerunds and Infinitives: The Case of EFL Learners

ZAINAB ALSUHAIBANI

Imam Mohammad Ibn Saud Islamic University, Saudi Arabia
alsuhaibani.z@gmail.com

ABSTRACT

The last decades witnessed the significant impact of using corpora on linguistics and language education. Particularly, the introduction of the direct application of corpora and Data-Driven Learning (DDL) has brought helpful tools in learning and teaching languages. This study investigated the effect of using corpora on teaching the confusing grammatical constructs, gerunds and infinitives, to EFL students. Further, the study explored EFL students' perceptions of using corpora in learning grammar. The study adopted a mixed method research design involving a triangulation of quantitative and qualitative data collection methods. A proficiency test and pre- and post-tests were administered to 89 sophomore EFL students. Further, a questionnaire and semi-structured interviews were used to explore EFL students' perceptions of using corpora in learning grammar. The results indicated a statistically significant difference between the control and experimental groups in favor for the group that received corpus-based instruction of gerunds and infinitives. In addition, the results showed students' positive perceptions of the use of corpora in learning grammar. The study concluded with some recommendations for ELT instructors and curriculum designers.

Keywords: grammar; EFL; DDL; corpora; corpus; gerunds; infinitives; instruction

INTRODUCTION

Since 1990s, the use of direct applications of corpora or what is called Data-Driven Learning (DDL) has gained considerable attention in language teaching and learning. Johansson (2007) commented that the advent of corpus linguistics has revolutionized research in linguistics and greatly influenced language learning and teaching. Despite the efforts and advances in corpus research, the use of corpora in teaching and learning has not been implemented greatly (Chambers, 2019; Poole, 2020). Granger (2015) maintained “the corpus world is replete with laments that the corpus revolution has not yet reached the language teaching world” (p. 507). As an EFL instructor attempting to bridge theory and practice, the researcher noticed the difficulties students find in learning English grammar, specifically with gerunds and infinitives. The difficulty lies in the need to memorize the verbs that can be used with gerunds or infinitives. This study attempts to help EFL learners in learning gerunds and infinitives through using corpora. It is hypothesized that the frequencies and the real authentic data found in corpora can give learners the opportunity to practice English gerunds and infinitives. Such continuous practice might help learners to cognitively process grammatical patterns to internalize them (O'keeffe et al., 2007).

Considering the literature of corpus research, most of the studies are about corpus-based analysis of gerunds and infinitives (Chalabian, 2020; Duffley & Fisher 2021; Kim & Yoo, 2015; Schwartz & Causarano, 2007). However, research on corpus-based instruction of gerunds and infinitives receives little, if no, attention. Further, the majority of corpus-based instruction studies investigate its effect on grammar, in general, rather than focusing on particular constructs, such as gerunds and infinitives (Abdul-Ameer,2019; Burton, 2022; Elsherbini & Ali, 2017; Lin, 2021).

More importantly, little attention has been paid to exploring students' perceptions of corpus-based instruction in learning grammar. Even when considering the few studies that explored students' perceptions of corpus-based instruction, there is a need for further studies as conflicting results have been reported (Lin, 2016; Lin, 2021).

Given such a gap in the literature along with the call for more innovative approaches for grammar instruction such as corpus-based instruction (Larsen-Freeman, 2015), the purpose of the study is to investigate the effect of using corpora on EFL learning of gerunds and infinitives. Additionally, the study aims to explore EFL students' perceptions of using corpora in learning grammar. More specifically, the study is set out to answer the following questions:

1. Is there any effect of corpus-based instruction on EFL learning of gerunds and infinitives?
2. What are EFL learners' perceptions of corpus-based instruction in learning grammar?

SIGNIFICANCE OF THE STUDY

This study is expected to have theoretical, methodological and practical significance. In theory, it is important to understand why and how using corpora can be beneficial to language learners, especially when considering EFL learning of gerunds and infinitives. Corpora offer learners real and authentic language; thus, they can help raise learners' awareness of how a particular language is used. The authentic language of corpora doesn't only show how a certain language is really used, but it also provides rich and more diversified information than dictionaries or reference grammar books (Braun, 2005). Through using corpora, not only teachers but also students can have immediate and free access to information about what can be considered acceptable or unacceptable in the language (Romer, 2006). It has also been argued that using corpora in L2 classrooms empowers learners' autonomy. It fosters students' role as active agents who are responsible for their learning, while considering the teacher as a mediator (McEnery and Wilson, 1997).

This study strives to investigate such theoretical assumptions while using a mixed method research design that combines both qualitative and quantitative measures in its investigation. Using a mixed method research design is especially important given the few studies that used such a design in investigating the effect of corpus-based instruction on grammar, in particular.

On practical grounds, the study is expected to provide insight into finding more effective ways for learning and teaching grammar. Specifically, it is hoped that the results of the study can raise teachers' awareness to consider corpus-based instruction for teaching grammar. Moreover, it is a desire that the current study may assist material designers in EFL contexts to consider designing grammar courses based on corpora.

LITERATURE REVIEW

CORPORA: DEFINITION AND CLASSIFICATION

A corpus is defined as "a collection of naturally occurring language texts, chosen to characterize a state or variety of a language" (Sinclair, 1991, p. 171). More specifically, Johnson and Johnson (1998) defined a corpus as "a large computer-held collection of texts (spoken, written or both) collected together to stand as a representative sample of a language or some part of it" (p.89-90). Broadly speaking, according to Römer (2011), corpora have direct and indirect pedagogical applications. Other researchers referred to them as 'hands-off' and 'hands-on' (Boulton, 2012).

Indirect pedagogical applications or hands-off applications might influence syllabus or materials design. In other words, researchers and material writers can use corpora to answer questions regarding what to teach and when to teach it. An example of such indirect pedagogical applications of corpora is the Collins COBUILD English course which has been developed depending on the frequencies of the common words in English (Romer, 2010). On the other hand, in the direct applications of corpora or the hands-on, learners and teachers are actively engaged in using corpora and concordances directly. Thus, they don't heavily depend on material designers for the provision of corpus-based materials. Tim Johns, the initiator of such an approach, suggested Data Driven Learning (DDL), in which the learner is faced as directly as possible with authentic data to explore linguistic patterns using concordances that help him become a linguistic researcher (Johns, 2002).

RATIONAL FOR DIRECT APPLICATION OF CORPORA

The direct application of corpora in language pedagogy has its own rationale and supporting arguments. Considering Schmidt's noticing hypothesis (1990-2001) and discovery learning that allow students to become language researchers (Gavioli, 2001; Johns, 2002), corpora can help learners notice certain structural patterns and discover how they behave in a given context. Further, Granath (2009) asserted that the most important advantage of using a corpus is the fact that it allows students to encounter 'real language' rather than abstract contrived examples (p.49). For example, it is better to have the learner confronted with real language written by native speakers in a newspaper instead of unauthentic examples. She also added that the authentic language of corpora provides students with a more vivid picture of the language than the one found in their reference books. Barnbrook (1996) maintained "students can derive the information they need directly from the language, as though the computer were a tireless native-speaker informant, with rather greater potential knowledge of the language than the average native speaker" (p.140).

In fact, using corpora in L2 classrooms promotes an autonomous and interactive kind of learning between students and language data (Pedrosa, Domínguez & Sintes, 2014). Bernardini (2002) described corpora as "rich sources of autonomous learning activities of a serendipitous kind" (p.165).

Moreover, using corpus-based instruction in language learning has positive effects on students' interest and motivation. Actually, learner autonomy and learner-centeredness, that are enhanced through the use of corpora, are conducive to increased motivation (Bernardini, 2002; Boulton, 2012; Lin, 2016; Liu, & Jiang, 2009). Aijmer (2009) commented that corpus-based language teaching can greatly motivate L2 learners.

RESEARCH ON DIRECT APPLICATION OF CORPORA

In spite of the benefits of using corpora in language learning, relatively few studies investigate the effect of the direct implementation of corpora in language classrooms (Boulton & Cobb, 2017; O'keeffe et al., 2007; Romer, 2011). The studies that explore the effectiveness of using corpora in teaching different areas of the language show positive results. More specifically, the direct application of corpora in language classrooms has actually been proven to be effective in teaching vocabulary (Binkai, 2012; Jafarpour et al., 2013; Jezo, 2013; Lee et al., 2019; Wu et al., 2019); grammar (Hunston, 2002; Lin, 2021; Liu & Jiang, 2009; Narita, 2012; Rodríguez-Fuentes, & Swatek, 2022); writing (Adel, 2010, Sun & Hu, 2020; Tsai, 2019; Yoon, 2008); and English for Specific Purposes (ESP) (Camiciottoli, 2010; Cheng, 2010; Hou, 2014).

Focusing on grammar, Hunston (2002) found that corpus-based teaching of grammar was greatly useful for promoting inductive learning and acquiring grammar by self-discovery which was, at the same time, entertaining and exciting. Oghigian and Chujo (2008) found that using corpus-based teaching of grammar with students was helpful in understanding word classes, derivations and in recognizing noun and verb phrases. Also, the study of Liu and Jiang (2009) revealed that integrating corpora and contextualized lexicogrammar in language teaching was effective in improving learners' command of lexicogrammar, increasing their critical understanding of grammar, and enhancing their discovery learning skills. Recently, Elsherbini and Ali (2017) found that learners who received corpus-based instruction outperformed the students who did not receive corpus-based instruction in learning grammar. In addition, Lin's study (2021) revealed that corpus-based instruction was effective in learning grammar for EFL university students regardless of their proficiency level.

Some studies focused on investigating the effect of corpus-based instruction on particular grammatical structures. For example, Ucar and Yukselir (2015) found that corpus-based instruction had a positive effect on learning verb-noun collocations in EFL classes. Further, Koosha and Jafarpour (2006) maintained that using Data-Driven Learning was highly effective in the teaching and learning of collocations of prepositions.

In sum, the results of these studies are promising in showing the positive effects of using corpora on learning grammar. However, there is a need to investigate the effect of corpus-based instruction on learning the confusing constructs of gerunds and infinitives, in particular. Further, there is a need to explore learners' perceptions of corpus-based instruction using a mixed-method research design. The following section sheds light on the how this study attempts to fill such a gap.

METHODOLOGY

SAMPLE

The sample of the study consisted of 89 sophomore EFL students majoring in English at a public university in Riyadh. Their ages ranged between 22 to 24. They had already passed the preparatory year and the first two levels in the department in which they took different courses of skills and subskills, including, reading, writing, listening, speaking, and grammar. Participants agreed to participate in the study filling out a consent form. They were informed about the study aim. They were also ensured that all their data will be kept confidential.

RESEARCH DESIGN AND INSTRUMENTS

A mixed-method research design incorporating both qualitative and quantitative data collection methods was followed to address the questions of the study. The use of both types of methods helps in getting a more in-depth analysis of the participants' perceptions of the use of a corpus in learning grammar, and the effect of using it on learning gerunds and infinitives. More specifically, the instruments of data collection were as follows:

PROFICIENCY TEST

A proficiency test adapted from TOEFL was used to ensure the homogeneity of the control and experimental groups before proceeding to the treatment of the study. TOEFL was chosen as it is

an international standardized test. The proficiency test contained four sections: reading, writing, grammar and vocabulary. Two complete hours were allocated for students to complete the test.

PRE-TEST AND POST-TEST

A pre-test and a post-test of gerunds and infinitives were used with both groups. The test contained two main fill-in-the-blank questions: a traditional fill-in-the-blank question with contrived examples (12 items), and a corpus-based fill-in-the blank question based on real data (12 items). Each of these two questions included equal structures of gerunds and infinitives. Gerund items included gerunds used after verbs, after adjectives, and as subject. Infinitive items also included items of the infinitive used after verbs, after adjectives, and as it-subject. The choice of the items went through a rigid process while considering two things. First, the frequencies of the gerund and infinitive structures in the two test questions were as close as possible. Second, the regularity of the verb forms used were carefully considered. The following table shows the chosen structures of gerunds and infinitives, and their frequencies in Corpus of Contemporary American English (COCA) :

TABLE 1. Gerund and infinitive structures used in pre- and post-test with their frequencies

	Structure	Traditional Fill-in-the blank	Frequencies	Corpus -based fill-in-the-blank	Frequencies
Gerund	After adjectives	Have fun playing	96	Busy working	128
		Interested in talking	145	Afraid of losing	217
	As subject	Learning	1878	keeping	1623
		sleeping	281	studying	384
	After verbs	Practice speaking	30	Finish reading	45
		Avoid making	241	Enjoy watching	192
Infinitive	After adjectives	Ready to take	1184	Happy to see	1250
		surprised to hear	521	Sorry to hear	432
	As it-subject	It is difficult to know	377	It is important to understand	311
		It takes time to walk	143	It costs to buy	219
	After verbs	Plan to travel	35	Agree to give	68
		Refuse to leave	200	Hope to find	203

QUESTIONNAIRE

The questionnaire was used to solicit information regarding the participants' perceptions of using corpora in learning grammar. The questionnaire consisted of two sections. The first section contains some demographic questions such as age and GPA (Grade Point Average). The second section includes twelve items that were adapted from different questionnaires in the literature (Elsherbini & Ali, 2017; Girgin, 2011). Thirty-one participants responded to these twelve closed-ended items using a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The reliability of the twelve items was measured using Cronbach's Alpha. The result indicates a high reliability coefficient (92).

TABLE 2. Questionnaire Reliability

Cronbach's Alpha	N of Items
.921	12

SEMI-STRUCTURED INTERVIEWS

Semi-structured interviews were also conducted to support the quantitative data obtained from the questionnaire. They mainly sought to explore the participants' perceptions of using the corpus in learning grammar, in general. They were also used to explore the difficulties the participants faced while using the corpus and whether they plan to use it in other courses or not. A total of nine students were interviewed. They were allowed to speak in Arabic (their native language) to help them become more comfortable and talk freely without being hindered by language proficiency level. They were asked four main questions with subsequent questions based on their answers. The main questions of the semi-structured interview were:

- What is your opinion about using the corpus for learning grammar?
- Did you face any difficulty in using the corpus?
- Are you going to use the corpus for your learning in future?
- Do you recommend using it in other courses?
- Is it better than the traditional way?

PROCEDURES

Different classes were randomly chosen and divided into control and experimental groups. A proficiency test was first used with both groups to ensure homogeneity. After piloting the test with a different sample and validating it from specialists in the field, the pre-test was also administered before beginning the study intervention. Then, the control group received explicit grammar instruction of gerunds and infinitives. The researcher explained the structures and uses of gerunds and infinitives found in the book *Grammar Sense 2* by Cheryl Pavlik, including uses of gerunds and infinitives after verbs, after adjectives, and as subject or it-subject. In the control group, students participated in class or in Blackboard discussion boards, did the exercises in the book, assignments, and worksheets based on the direct explanation of gerunds and infinitives.

On the other hand, the experimental group received direct corpus-based instruction of gerunds and infinitives using the Corpus of Contemporary American English (COCA). COCA was chosen as it is the most widely used corpus in English which contains more than one billion words from different genres. More importantly, it allows free access in a clear interface which helps students to use it free from any charge. The researcher first introduced the structures and uses of gerunds and infinitives using the same book and explained gerunds and infinitives directly using the corpus. Meanwhile, she trained the students in the experimental group on how to register and use the corpus. She also gave them a brochure to guide them in the registration and looking for examples. Two hours were allocated for training students. Students in the experimental group were asked to use the corpus in all class activities, including participation in class or in Blackboard discussion boards, and in answering the exercises and worksheets.

After the intervention which lasted six hours in three weeks (two hours a week), the researcher administered the post-test to both groups. The researcher then distributed the questionnaire and conducted semi-structured interviews with students from the experimental group who voluntarily agreed to participate. To analyze the data, the pre- and post-tests were corrected considering one mark for each correct use of gerunds or infinitives. SPSS was used to analyze the difference between the means of the control and experimental groups. For the questionnaire, the means and standard deviations for the items were calculated. Content analysis was used for analyzing the qualitative data.

RESULTS

THE EFFECT OF USING CORPUS-BASED INSTRUCTION ON EFL LEARNING OF GERUNDS AND INFINITIVES

SPSS was used to analyze the quantitative data. First, independent sample t-test was used to explore whether there are any differences between the control and the experimental groups in their proficiency test or GPA scores. The means and standard deviations were calculated as shown in Table 3. The results of the tests (Table 4) revealed that there are no significant differences between the control and experimental groups in terms of proficiency test $t(87)=.084$, $p=.933$, and GPA scores $t(87)=.139$, $p=.890$ indicating the homogeneity of the two groups.

TABLE 3. Means and standard deviations of the proficiency and GPA scores

	Group	N	Mean	Std. Deviation
Proficiency test	control	39	25.2179	10.69050
	experimental	50	25.0200	11.31820
GPA	control	39	3.7605	.81928
	experimental	50	3.7386	.67300

TABLE 4. Independent sample t-tests for the proficiency test and GPA scores

		Levene's Test for Equality of Variances		t-test for Equality of Means			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
proficiency test	Equal variances assumed	.064	.801	.084	87	.933	.19795
	Equal variances not assumed						
GPA	Equal variances assumed	2.856	.095	.139	87	.890	.02191
	Equal variances not assumed						

Further, the results in Table 6 showed that there are no significant differences between the control group ($M= 11.82$, $SD= 5.09$) and the experimental group ($M=11.86$, $SD= 6.64$) in the pre-test of gerunds and infinitives $t(87)= .032$, $p=.975$. Yet, statistically significant differences were found between the two groups in the post-test $t(87)=. 3.232$, $p = .002$. Such results indicate that

there is an effect of corpus-based instruction on learning gerunds and infinitives. To measure the effect size of corpus-based instruction on learning gerunds and infinitives, Hedges' *g* was calculated given the different sample sizes and standard deviations of the two groups. The results revealed a medium effect size of corpus-based instruction of gerunds and infinitives (0.69) which indicates its effectiveness in this study.

TABLE 5. Means and standard deviations in the pre-test and post-test

	Group	N	Mean	Std. Deviation
Pre-test	control	39	11.8205	5.09836
	experimental	50	11.8600	6.64250
Post-test	control	39	16.5128	5.53872
	experimental	50	19.8800	4.29352

TABLE 6. Differences between groups in the pre-test and post-test

		Levene's Test for Equality of Variances		t-test for Equality of Means			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Pre-test	Equal variances assumed	4.686	.033	-.031	87	.976	-.03949
	Equal variances not assumed			-.032	86.985	.975	-.03949
Post-test	Equal variances assumed	3.867	.052	-3.232	87	.002	-3.36718
	Equal variances not assumed			-3.133	70.038	.003	-3.36718

Even when considering only the traditional fill-in-the-blank question in the pre- and post-tests to avoid the priming effect of the corpus, the difference between the control ($M= 8.05$, $SD= 2.90$) and experimental ($M= 9.80$, $SD= 2.21$) groups is statistically significant $t(87)= 3.22$, $p = .002$ as shown in Table 7 and 8.

TABLE 7. Means and standard deviations in the pre-test and post-test of the traditional fill-in-the-blank question

	group	N	Mean	Std. Deviation	Std. Error Mean
Pre-test	control	39	6.0513	2.79989	.44834
	experimental	50	5.7800	3.39441	.48004
Post-test	control	39	8.0513	2.90144	.46460
	experimental	50	9.8000	2.21313	.31298

TABLE 8. Differences between groups in the pre-test and post- test in the traditional fill-in-the-blank question

		Levene's Test for Equality of Variances		t-test for Equality of Means			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Pre-test	Equal variances assumed	.890	.348	.403	87	.688	.27128
Post-test	Equal variances assumed	3.019	.086	-3.227	87	.002	-1.74872

EFL LEARNERS' PERCEPTIONS OF USING CORPUS-BASED INSTRUCTION IN LEARNING GRAMMAR

Considering students' perceptions of using corpus-based instruction in learning grammar, the results of the questionnaire in Table 9 indicate that EFL students have positive perceptions of using the corpus in learning grammar.

TABLE 9. Means and standard deviations of the perception questionnaire

Statement	Mean	Std. Deviation
Using the corpus for learning grammar is easy.	3.97	.948
Using the corpus for learning grammar is useful.	4.10	.944
Using the corpus for learning grammar is interesting.	3.94	.998
Using the corpus for learning grammar is more helpful than a dictionary	3.87	1.118
Using the corpus for learning grammar increases my confidence.	3.90	1.012
Using the corpus for learning grammar learning helps me remember the rule better.	4.03	.912
Learning grammar through the corpus is better than traditional methods.	3.71	1.131
I enjoy discovering grammar rules through the corpus.	3.84	.934
I can see the target sentences in real use through the corpus.	4.23	.845
I like using the corpus for learning grammar.	3.87	.846
I develop my knowledge in grammar through using the corpus.	4.00	.894
I recommend using the corpus in other courses.	3.97	1.110

The qualitative data obtained from students' responses to the interview questions also supports that most students find the corpus useful. When asked about their opinion of the corpus, most of the students mentioned that it is helpful to learn grammar, interesting, important, rich with examples, innovative, etc. One student stated, "it is very good. It helped me understand some of the points that I find difficult." Another maintained "it saves time with the authentic examples." Further, one elaborated that it is useful in learning and better than traditional ways. One student also explained that whenever she feels confused about the use of a structure, she just looks for the word in the corpus and gets many examples that help her. Only one student said that she doesn't like the corpus.

When asked about the difficulties in using the corpus for learning grammar, students' answers vary. Some maintained that it was easy, clear and that they didn't encounter any difficulties. However, other students stated that they faced difficulties, especially in the registration at the beginning. One student maintained "I faced difficulties in the registration at the beginning. After that, everything was easy and clear in the website interface." Another stated that she

sometimes faced difficulties as she might have not understood how to use it following the instructions. In addition, two students mentioned that they found difficulties in looking for clear examples. Only one student complained that she found it sometimes hard to use and slow with the corpus freezing. Generally, the students maintained that it was difficult at the beginning. But after following the instructions and after getting used to it, it became easy and clear: “it is good and easy once we get familiar with it”, one student stated.

Most of the students also recommended using it in other courses and mentioned that they are going to use it in their coming courses. One student said “Sure, I will use it as it is important to understand meaning and build correct structure.” Another student maintained that she has already begun to use it in other courses like writing. Another student has sent the website link to her colleagues in the writing class, as well. One student explained that the corpus should not only be used in grammar, but also in developing English language in general. The interview also showed that the students use the corpus for collocation, modeling examples to build correct sentences, and for checking errors.

DISCUSSION

The results of the study indicated that corpus-based instruction is effective in learning gerunds and infinitives. Such results mirror the findings of previous studies on the effect of the use of corpora on learning grammar in general (Abdul-Ameer, 2019; Elsherbini & Ali, 2017; Hunston, 2002; Oghigian & Chujo, 2008; Lin, 2021; Liu & Jiang, 2009; Narita, 2012), and in learning specific grammatical structures, in particular (Jafarpour, 2006; Rodríguez-Fuentes, & Swatek, 2022; Ucar & Yukselir, 2015). For example, considering studies of learning grammar in general, Lin (2021) investigated the effect of using corpora on enhancing EFL students’ grammatical competence with different levels of proficiency. Using pre- and post-tests, he found that students from all proficiency levels in the experimental group outperformed the students in the control group in grammar. Even when only considering corpus-informed materials, the results are promising. For gerunds and infinitives, for example, Rodríguez-Fuentes and Swatek (2022) investigated the effect of corpus-informed textbooks on learning gerunds and infinitive among 89 EFL learners in Colombia. Although both treatment and control groups showed improvement; yet the results of the pre-test, post-test, and delayed test all indicated the superiority of the group who received corpus-informed materials. It seems that using corpora with the feature of frequencies help learners consciously ‘notice’ the differences in the use of gerunds and infinitives which can raise their awareness and help to develop the input into intake (Schmidt, 1990, 2001). More importantly, the practice of consulting examples in corpora allows learners to be confronted with authentic language that helps them in the process of self-discovery and remembering.

In addition, the results showed EFL students’ positive perceptions of using corpus-based instruction for learning grammar. This is also supported by many studies that explored learners’ perceptions and attitudes towards the use of corpora in grammar learning (Elsherbini & Ali, 2017; Kim, 2019; Lakew et al., 2021; Lin, 2016). In the current study, students found using the corpus helpful, interesting, and an innovative way of learning grammar. They also recommended using it in other courses. In fact, students don’t have only positive perceptions of using the corpus in grammar, but also corpus-based instruction can help students enhance their motivation and self-efficacy in learning grammar (Lin, 2016).

It is true that students have positive perceptions of using the corpus for learning grammar, yet some expressed certain difficulties they have encountered in using it, such as problems in registration and understanding examples. Similarly, Elsherbini and Ali (2017) found that students faced technical problems in registration to COCA and interpreting examples of concordances. In fact, as a new technological tool, it is expected that students will face difficulties. This is accompanied by the fact that the free access requires registration. For that reason, students need to be first trained on how to register and use the corpus with focus in a particular word/structure in order not to be lost with data. This has also been reflected in the interview results as students in general maintained that such difficulties were mostly at the beginning, and they could overcome them with practice.

CONCLUSION

The study finds empirical evidence supporting the use of corpus-based instruction in EFL learning of English gerunds and infinitives. It also shows EFL students' positive perceptions regarding the use of corpora in learning grammar. Thus, it is recommended to implement corpus-based instruction in EFL setting. Using corpora in EFL setting is particularly important as they can compensate for the lack of authentic materials. Yet, teachers need to be trained first on how to use corpora in language classrooms while considering the constraint of time and curriculum. More importantly, students' level and age should be taken into consideration before using corpora directly. Further, training on the use of corpora should precede students' direct use to ensure effectiveness. Also, guidelines in registration and searching through the concordance lines need to be available to students to ensure practicality.

Moreover, indirect applications of corpora are needed. It is recommended that material designers build their materials considering the authentic examples found in corpora. The provision of frequency tables can inform learners about different aspects of vocabulary and structure, such as collocations and correct use. All in all, corpora can be of help to language learners, especially in the EFL setting.

However, there are several limitations in this study that should be taken into considerations. First, the study focused on only corpus-based instruction of gerunds and infinitives. Other structures need to be studied in future research. Further, the study did not investigate the long-term effect of corpus-based instruction on learning gerunds and infinitives. Thus, studies with longer intervention and delayed post-tests are needed.

ACKNOWLEDGMENT

This research was supported by the Deanship of Scientific Research, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia, Grant No. (20-13-07-002).

REFERENCES

- Abdul-Ameer, M. A. (2019). Corpora in the EFL classroom: Exploring the effects of data driven learning (DDL) on Iraqi EFL freshmen's grammatical development. *Journal of Basra researches for Human Sciences*, 44(4-B).
- Adel, A. (2010). Using corpora to teach academic writing: Challenges for the direct approach. In M. Campoy-Cubillo, B. Belles-Fortuño & L. Gea-Valor (eds.), *Corpus-based approaches to English language teaching* (pp. 39-55). London: Continuum.

- Aijmer, K. (Ed.). (2009). *Corpora and language teaching*. Amsterdam: John Benjamins
- Barnbrook, G. (1996). *Language and computers: A practical introduction to the computer analysis of language*. Edinburgh: Edinburgh University Press
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165-182). Amsterdam: Rodopi.
- Binkai, J. (2012) An empirical study on corpus-driven English vocabulary learning in China. *English Language Teaching*, 5 (4) 131-137.
- Boulton, A. (2012) Language awareness and medium-term benefits of corpus Consultation. In Gimeno Sanz, A. (ed.), *New Trends in Computer-Assisted Language Learning: Working Together* (pp. 39-46.). Madrid: Macmillan.
- Boulton, A. (2012). Hands-on/hands-off: Alternative approaches to data-driven learning. In *Input, process and product: Developments in teaching and language corpora* (pp. 152-168). Masarykova univerzita.
- Boulton, A. (2012). Hands-on / hands-off: Alternative approaches to data-driven learning. In J. Thomas & A. Boulton (eds), *Input, Process and Product: Developments in Teaching and Language Corpora*. Bmo: Masaryk University Press, p. 152-168
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(23), 48–393.
- Braun, S. (2005) From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, 17(1): 47-64.
- Burton, G. (2022). What can a corpus tell us about grammar teaching materials?. In *The Routledge Handbook of Corpus Linguistics* (pp. 358-370). Routledge
- Camiciottoli, B. (2010) Corpus-informed approach to teaching lecture comprehension skills in English for business studies. In M. Campoy-Cubillo, B. Belles-Fortuño & L. Gea-Valor (eds.). *Corpus-based Approaches to English Language Teaching* (pp. 95-107). London: Continuum.
- Chalabian, F. (2020). *ESL textbooks materials and real language use: Comparing corpus-based materials and textbook materials on gerunds/infinitives* (Doctoral dissertation, Carleton University).
- Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching*, 52(4), 460-475.
- Cheng, W. (2010). Hong Kong engineering corpus: Empowering professionalism-training to learn the language of their profession. In M. Campoy-Cubillo, B. Belles-Fortuño & L. Gea-Valor (eds.). *Corpus-based Approaches to English Language Teaching* (pp. 67-79). London: Continuum.
- Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.
- Duffley, P., & Fisher, R. (2021). To-Infinitive and Gerund-Participle Clauses with the Verbs Dread and Fear. *Studia Linguistica*, 75(1), 72-96.
- Elsherbini, S. A. H., & Ali, A. D. (2017). The effects of corpus-based activities on EFL university students' grammar and vocabulary and their attitudes toward corpus. *Journal of Research in Curriculum Instruction and Educational Technology*, 3(1), 133-161.
- Gavioli, L. (2001). The learner as researcher: Introducing corpus concordancing in the classroom. In Aston, G. (Ed.), *Learning with corpora* (pp. 1000–1130). Athelstan/Bologna.
- Girgin, U. (2011). *Corpus-based activities at lower levels of EFL proficiency: The effectiveness of using concordance lines on grammar learning* (Master's thesis, Eğitim Bilimleri Enstitüsü).
- Granath, S. (2009). Who benefits from learning how to use corpora? In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 47-65). Amsterdam, Netherlands: John Benjamins.
- Granger, S. (2015). The contribution of learner corpora to reference and instructional materials design. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The cambridge hand- book of learner corpus research* (pp. 486–510). Cambridge: Cambridge University Press.
- Hou, H. (2014) Teaching specialized vocabulary by integrating a corpus-based approach: Implications for ESP course design at the university level. *English Language Teaching*, 7(5), 26-37
- Hunston, S (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Jafarpour, A. A., Hashemian, M., & Alipour, S. (2013). A Corpus-based Approach toward Teaching Collocation of Synonyms. *Theory and Practice in Language Studies*, 3(1), 51-60.
- Johansson, S. (2007), Using corpora: From learning to research. In: Hidalgo, E., Quereda, L., Santana, J. (eds). *Corpora in the foreign language classroom* (pp.17-30). Amestrdam: Rodopi.
- Johns, T. (2002). Data-driven learning: The perpetual challenge. *Language and Computers*, 42(1), 107-117.

- Johnson, K. & Johnson, H. (eds.) (1998) *Encyclopedic dictionary of applied linguistics*. Massachusetts: Blackwell Publishers Ltd.
- Kim, H. (2019). The perception of teachers and learners towards an exploratory corpus-based grammar instruction in a Korean EFL primary school context. *Primary English Education*, 25(1), 123-152.
- Kim, J. E., & Yoo, I. W. (2015). A Corpus-based Study of To-Infinitive Errors in Korean College Freshmen's Writing. *Journal of Asia TEFL*, 12(4).
- Koosha, M., & Jafarpour, A. A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL journal*, 8(4), 192-209.
- Lakew, A. K., Teshome, S., & Negede, D. (2021). Introducing Corpus-Informed Spoken Grammar Instruction in EFL Classrooms: The Ethiopian Experience. *Corpus Pragmatics*, 5(4), 487-510
- Larsen-Freeman, D. (2015). Research into practice: Grammar learning and teaching. *Language Teaching*, 48(2), 263-280.
- Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, 40(5), 721-753.
- Lin, M. H. (2016). Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *Tesol Quarterly*, 50(4), 871-893.
- Lin, M. H. (2021). Effects of data-driven learning on college students of different grammar proficiencies: A preliminary empirical assessment in EFL classes. *SAGE Open*, 11(3), 2158244021102993
- Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, 93 (1), 61-78.
- McEnery, T. & A. Wilson (1997) Teaching and language corpora. *ReCALL*, 9(1), 5-14.
- Narita, M. (2012) Developing a corpus-based online grammar tutorial prototype. *The Language Teacher*, 36 (5), 23-29.
- O'Keefe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Oghigian, K., & Chujo, K. (2010). An effective way to use corpus exercises to learn grammar basics in English. *Language Education in Asia*, 1(1), 200-214.
- Pedrosa, A. V., Domínguez, J., & Sintés, A. (2014). Introduction: the use of corpora for language teaching and learning. *Research in Corpus Linguistics*, 1(1), 1-5.
- Poole, R. (2020). "Corpus can be tricky": Revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 1-22.
- Rodríguez-Fuentes, R. A., & Swatek, A. M. (2022). Exploring the effect of corpus-informed and conventional homework materials on fostering EFL students' grammatical construction learning. *System*, 104, 102676.
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205-225.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Schmidt, R.W. (2001) Attention. In Robinson, P. (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schwartz, M., & Causarano, P. N. L. (2007). The role of frequency in SLA: An analysis of gerunds and infinitives in ESL written discourse. *Journal of Second Language Acquisition and Teaching*, 14, 43-57.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press
- Sun, X., Hu, G. (2020). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research*. <https://doi.org/10.1177/1362168820954459>
- Tsai, Y. R. (2021). Exploring the effects of corpus-based business English writing instruction on EFL learners' writing proficiency and perception. *Journal of Computing in Higher Education*, 33(2), 475-498.
- 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.
- Wu, S., Fitzgerald, A., Yu, A., & Witten, I. (2019). Developing and evaluating a learner-friendly collocation system with user query data. *International Journal of Computer-Assisted Language Learning and Teaching*, 9(2), 53-78.
- Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning & Technology*, 12(2), 31-48.