

Saudi students' attitudes towards the use of corpora in learning collocations

العنوان: اتجاهات الطلاب السعوديين نحو استخدام المتون اللغوية في تعلم المتلازمات اللفظية

عواطف كاتب سبيعي الرويلي

الملخص: استخدمت لغويات المتون الحاسوبية لأكثر من ثلاثة عقود في تدريس اللغة الإنجليزية لكن إلى الآن لم يصبح منهجا سائداً لتعلم اللغة في الفصول الدراسية في المملكة العربية السعودية. هناك نقص في الدراسات المنشورة التي تبحث استخدام لغويات المتون الحاسوبية في تدريس اللغة الإنجليزية في السياق السعودي. ولمعالجة هذه الفجوة، قيمت هذه الدراسة استخدام لغويات المتون في الفصول الدراسية للغات الأجنبية (اللغة الإنجليزية) عن طريق دراسة مجموعة من الطلاب في جامعة الحدود الشمالية في المملكة العربية السعودية. على وجه التحديد، بحثت هذه الدراسة مواقف الطالبات نحو نهج التعلم القائم على البيانات (DDL) بشكل عام، واستخدام مورد (AntConc)، ومواد التعليم القائمة على concordance-based بشكل خاص والتي تم استنباطها باستخدام الاستبيانات والمقابلات المتعمقة.

تشير البيانات النوعية والكمية للدراسة إلى أن الطالبات بوجه عام لديهن مواقف إيجابية تجاه نهج DDL ويرون أنه مفيد لتطوير اللغة الإنجليزية، وزيادة الثقة والكفاءة في لغتهم الإنجليزية. تشير النتائج إلى أن الطالبات كانوا قادرين على تقديم تعميمات وتكيفات مفيدة لنهج DDL في التعامل مع خطوط التوافق. كانوا أيضاً إيجابية تجاه نهج DDL وكانوا قادرين على التعرف على فوائد لاستخدامها في التعلم المتلازمات اللفظية (الفعل - الاسماء).

Title: Saudi students' attitudes towards the use of corpora in learning collocations

Awatif Katib Alruwaili

College of Education and Arts, Northern Border University, Arar, Saudi Arabia

Abstract

Corpus linguistics has been used for over three decades in language teaching but not until now has it become a mainstream approach to language learning in the classroom in Saudi Arabia. There is a lack of published studies that explore the use of corpora in teaching English in the Saudi context. To address this gap, the study evaluated the use of corpora in English foreign language (EFL) classrooms by examining a group of students at a Northern Border University in Saudi Arabia. Specifically, this study examined students' attitudes towards the data-driven learning approach (DDL) in general, the use of a corpus resource (AntConc), and concordance-based materials in particular that were elicited using reflective forms and in-depth interviews.

The study's qualitative and quantitative data indicate that overall, the students have positive attitudes towards the DDL approach and see it as beneficial to developing the English language; and increased confidence and proficiency of their English language. The results indicate that the students were able to make useful generalisations and adaptations of the DDL approach in dealing with concordance lines. They were also positive towards the DDL approach and were able to recognise the benefits for using it in learning collocations (verb–noun collocations).

Keywords:

Authentic materials, attitudes, collocations, corpora, data-driven learning.

1. Introduction

In terms of its application as a resource in language education, corpus-based learning has gained interest because of its ability to facilitate language acquisition for learners (Bernardini, 2004). However, it was only recently that a corpus as a pedagogical resource received attention within second language acquisition (SLA); as Fox stated, ‘the use of concordances in the classrooms is in its infancy as a language teaching technique’ (1998, p. 43). Yet corpus linguistics (CL) has not entered the mainstream of language teaching (e.g., Frankenberg-Garcia, 2012), and ‘the direct uses of corpora in language teaching are treated rather marginally in the literature in the field’ (Leńko-Szymańska & Boulton, 2015, p. 3).

More specifically, implementing the DDL approach in the Saudi context seems an even more distant possibility, due in part to the lack of published studies that explore the use of corpora in teaching English in the Saudi context. Alongside this lack of research, there is also a lack of knowledge of CL despite encouragement by the Ministry of Higher Education to develop and improve the English curriculum and teaching methodologies in Saudi Arabia. Based on the aforementioned problems, the present study aims to investigate the students’ attitudes towards the use of the DDL approach in the language classroom.

2. Corpus linguistics and its contributions to second language pedagogy

One major contribution of CL is its data-driven and empirical description of language, which is based on language in use, whether written or spoken (Biber et al., 1999), in turn enriching knowledge of that language. This unbiased description of the language has radically changed the design of dictionaries, starting with the influential COBUILD project at Birmingham University (Sinclair, 1987), which has been followed by many famous publishers of pedagogical materials (Cobb & Boulton, 2015). This influence on material design has been extended to include grammar books, such as the *Longman Grammar of Spoken and Written English* (Biber et al., 1999); these reference books are

influential because they were designed based on ‘real usage and frequency data’ (Cobb & Boulton, 2015, p. 479). Here, real usage can be summarised by providing more reliable information on how different aspects of language are used, such as judgement on collocations, frequency, semantic prosody and pragmatic meaning (Hunston, 2002).

There is a consensus that CL is a valuable source of information on actual language use (Chambers & O’Sullivan, 2004). Gavioli (2005) noted that concordance lines can enhance a teacher’s ability to gauge difficult or problematic areas, such as synonyms. The use of corpora in teaching and learning can be fruitful because it provides an opportunity for students to interact with real language, allowing teachers to manipulate this real language in various ways—for example, to illustrate a variety of registers (Granath, 2009). Learner autonomy is also supported by the independent use of corpora (Chambers, 2005). In addition, corpora may increase teachers’ and learners’ language awareness and hence facilitate the development of the learning skills (Breyer, 2009; Farr, 2008) by providing information related to frequency, distribution across registers and patterns of target items.

In this regard, it is important to mention that the early use of corpora in the classroom was in the form of the DDL approach (Johns, 1991a), which is based on a communicative method that focuses on teaching languages according to communicative competence rather than performance (Richards & Rodgers, 2014). The DDL approach is known for encouraging a student-centred learning process and fostering learner autonomy. In addition, Cheng (2012) asserted that this approach can be used to improve explicit and implicit learning, which eventually leads to the improvement of the explicit and implicit acquisition of knowledge. Given the complexity of the DDL approach and its use in the classroom, the following section elaborates on the DDL approach and the reality of its implementation in the EFL context.

2.1. Data-driven learning approach

The DDL approach was first introduced to the field of second language (L2) learning (SL) by Johns (1986). The DDL approach emerged as a result of the COBUILD project, the goal of which was to teach English by exploiting authentic language. The DDL approach can be defined as ‘the use in the classroom of computer-generated concordances to get students to explore the regularities of patterning in the target language’ (Johns & King, 1991, p. iii). At the beginning of his project, Johns prepared teaching materials using a concordance from the Bank of English (a corpus consisting of 4.5 million words of English texts) and printed concordance lists out as hard copy handouts to be used as exercises in the classroom, instead of using an online concordancer. However, with the development of computer software, the use of an online corpus in teaching has become easier. To conclude his project, Johns (2002, p. 109) stated that the DDL approach is more ‘effective’ in teaching lexis–grammar interfaces, especially when it comes to collocations. The DDL approach is based on an inductive approach towards teaching (Bernardini, 2004), where the students analyse the targeted features with or without the teachers’ help. After using it for 3 to 4 years with international students at the University of Birmingham, Johns (1991b, p. 2) reached important conclusions on the effective role of the DDL approach in the language learning process: The DDL approach is useful in developing the learners’ abilities to discover patterns and make generalisations, and it encourages inductive learning by giving the learners the opportunities to search for the rules through the authentic examples. However, what distinguishes the DDL approach from other methods of inductive learning is that the teacher does not ‘know exactly’ what rules the learners may grasp from those concordance lines (Johns, 1991b, p. 3). Also, the teacher’s role changes from a traditional one to a director or coordinator of the learners’ research. However, this change could be difficult for some teachers, especially when learners ask difficult questions or questions to which they have not prepared answers. In addition, the most important effect is that the DDL approach is ‘a reevaluation of the place of grammar in language learning and

language teaching' (Johns, 1991b, p. 3); it tackles all grammar rules that have not been taught before because they were 'too difficult or....overlooked'. The DDL approach draws the attention of learners towards grammar discovery and considers this discovery to be the 'centre of language learning' (Johns, 1991b, p. 3). This discovery is derived from the authentic examples to which the learners are exposed.

Sinclair (1999) noted that corpus data can enhance language teaching and that the form–meaning link can be taught through these data to minimise the learning load. Thus, learners can use the corpus evidence to develop their creative use of language or 'creative processing' (Nation, 2013, p. 110) that occurs when the learners come across target items that they already knew but that have different meanings and/or in different contexts. Coming across these target items encourages learners to 'reconceptualise their knowledge' of these items (Nation, 2013, p. 110). Chan and Liou (2005) revealed that delexicalised verbs are more effectively learnt through web-based concordancing, and Cobb's (1997) findings showed that learners are able to acquire knowledge of new vocabulary through concordances.

The DDL approach is repeatedly mentioned in the literature as being beneficial in L2 learning (Pérez-Paredes et al., 2011; Vyatkina, 2016a, 2016b; Smart, 2014) because of the rich input in the form of authentic data. However, there are some controversial issues that have gained interest in the literature: learner's proficiency, training and guided use. Yoon and Hirvela (2004) showed that the learners' levels do not have a significant effect on the learners' perceptions and attitudes towards the use of corpus, as the lack of familiarity and training can do. Most studies (i.e., Sun & Wang, 2003; Chan & Liou, 2005; Cresswell, 2007; Boulton, 2010c; Pérez-Paredes et al., 2011; Vyatkina, 2016a, 2016b; Smart, 2014) have demonstrated that the gradual and guided use of the DDL approach is more effective than an unguided one. However, the results have differed in various contexts. For example, the two studies by Cresswell (2007) and Chan and Liou (2005) were conducted in a similar environment (FL context), yet each study had a different focus, worked with different proficiency levels and

targeted different foci. Unsurprisingly, each study's results were different. Cresswell's (2007) study found that the treatment was insignificant and that the learners did not benefit from using the DDL approach in the classroom, whereas Chan and Liou's (2005) study indicated that the treatment was significant regarding some types of collocations, and the learners benefited from using a concordancer. As discussed on many occasions, it is quite rare to find studies that would provide hands-on training and investigate the learners' actual use of corpora at the same time. This is why the present research provides hands-on training and guided use before eliciting the participants' attitudes.

There are studies that investigate the uses of corpora in the classroom, and these studies have focused on the learners' attitudes and perceptions (e.g., Hafner & Candlin, 2007; Yoon & Hirvela, 2004; Chambers & O'Sullivan, 2004; O'Sullivan & Chambers, 2006; Varley, 2009; Geluso & Yamaguchi, 2014). Although these studies shared the same focus, they were different at many levels: the context of the study, the use of corpus applications in the classroom, the learners' proficiency and the targeted skill. For example, Hafner and Candlin (2007), Yoon and Hirvela (2004), Chambers and O'Sullivan (2004) and O'Sullivan and Chambers (2006) focused on improving the writing skills of learners, whereas Varley (2009) was more concerned with the grammar, and Geluso and Yamaguchi (2014) focused on improving the learners' fluency. The above-mentioned studies indicate that training learners in the use of the DDL approach and in using corpus tools is essential to fostering positive attitudes towards the corpus tools and the DDL approach. At the same time, it seems that learners' proficiency levels do not have a significant effect on learners' attitudes, which the present research shows as well. Through reviewing all these studies, none of them targeted the Saudi context or provided details on the training. In addition, none have focused on teaching collocations (verb-noun) when using the DDL approach. Thus, the current research is significant for the FL context in general and the Saudi context in particular.

3. The study

3.1. Research questions

The present study addresses the following four research questions:

- 1. Do the participants' ability to use computers have any effect on their use of the corpus resource?**
- 2. What are the participants' overall evaluations of the use of the DDL approach?**
- 3. What difficulties do the participants have in using a corpus?**
- 4. Did the training have an effect on the participants' attitudes? If so, how?**

3.2. Participants and context of the study

The students involved in this research were female students from Northern Border University (NBU) in Saudi Arabia; these participants were Arabic-speaking undergraduates and participated in the current study in exchange for course credit. Fifty-one students were enrolled in the study who each completed three training sessions and two testing sessions.

According to the university placement test, the participants were intermediate-level EFL students, in their foundation year and majoring in one of four medical fields: medicine (N=21), laboratory medicine and pathology (N=13), nursing (N=11) and pharmacy (N=6). According to my experience and familiarity with the students, I would suggest their level to be B1, which would be equal to a band score of 4–4.5 on the IELTS test. However, according to the self-report questionnaire, 19.6% rated themselves as beginner, 58.8% as lower-intermediate and only 21.6% as intermediate level. The participants were between the ages of 18 and 25, and they had studied English in school for approximately seven years. I was unable to obtain official proficiency measures for the participants, but all the participants were successfully enrolled in university-level English classes. They were divided into two groups because of the lab capacity, with 25 in the first group and 26 in the second group. The study ran for six weeks for one

group; the second training session was divided into two short sessions due to a scheduling conflict, which meant that the participants could not attend the sessions for more than 30 minutes each week. For the other group, the study took five weeks as planned. These sessions were held for one hour each week as part of the English curriculum's extracurricular activities. The participants attended 15 hours/week and followed the *New Headway* series.

Six participants volunteered to take part in the interview part of the study. I used a variation sampling strategy for selecting the participants, in which the participants differed in terms of their majors and proficiency levels, as well as whether they liked and/or disliked the experiment. Thus, when I recruited the participants, I went to each class and asked for volunteers from different majors and whether they liked or disliked the experiment, and then, interviewed all the participants who volunteered. Four students rated their proficiency at a low-intermediate level; here, three were nursing students, and one majored in pathology. The two remaining participants rated their proficiency as intermediate level; here, one was a medical student, and the other was a pharmacy student. Each of these students volunteered to take part in the recorded interview.

3.3. Methodology

The present study was divided into two sections: training sessions and testing sessions. Each session was held in a computer laboratory for one hour. The training sessions focused on how to use the corpus tool, AntConc (specifically, its concordancer). The three training sessions included a short introduction to CL and how it is used in language analysis. During the training sessions, the participants were trained to use AntConc to investigate verb-noun collocations following Sinclair's model (2003). The targeted collocations called general verbs (GVs) in which the verb carries a general semantic meaning, whereas the noun denotes a specific meaning such as *make a mistake/ sense*. Sinclair's (2003) framework informed the design for the teaching materials used in the current experiment, and the activities were divided according to this framework, which consists of seven steps: initiate,

interpret, consolidate, report, recycle, result and repeat (Sinclair, 2003, pp. xvi-xvii). This model seems to be designed for use by researchers more than by learners because it is time-consuming, which may demotivate learners based on the number of steps appearing in each session. Thus, I modified the design slightly so that it could be used in the classroom, with the revised model consisting of the following components: initiate, interpret, consolidate and report. During the testing sessions, the participants were asked to perform tasks similar to the training sessions but with different items, thus following the same procedure set in the training sessions. The reflective form consisted of three sections: a) background information (country of origin, major, age, number of years studied, proficiency, computer literacy and familiarity with the corpus tools); b) evaluations on using the concordancer for the tasks, such as how easy the corpus was to use, how helpful the approach they followed in reading the concordance lines was, and the advantages and potential disadvantages of using the DDL approach for learning language; c) the future intentions of using the corpus tool after the current study. The reflective form contained open-ended questions that comprised a qualitative data source. The reflective form used in the present study was designed based on the forms used in Varley's study (2009) but with modifications.

The second method used to elicit the students' evaluations was the interview, which aimed to provide an overall evaluation of the use of a corpus in the classroom. The interviews were semistructured, offering a space for the participants to further comment on their evaluations and attitudes towards the DDL approach and tool. In-depth interviews on the topic presented in the questionnaire were also carried out with six students to obtain more details. The interviews relied on a set of prepared questions yet were flexible in adding follow-up questions to elaborate on the answers of the participants as and when necessary. The interview questions were guided by the reflective form's themes, according to Varely's article (2009). The interview guide consisted of initial questions about computer literacy and then addressed the students' experiences and attitudes towards the DDL approach, the tasks

and the concordancer tool. All interviews were conducted in Arabic and audio-recorded using a digital recorder. The interviews (Arabic) were transcribed verbatim. Willig (2013, p. 31) said, 'It is sufficient to transcribe what is being said (the words)' if the researcher's aim was the content. Thus, the nonlinguistics features of the speech were excluded, and the content (the words) only was transcribed. Richards (2003) advised to create a set of transcription symbols before starting transcribing or at least when transcribing the first interview, which will help in holding a systematic approach for the transcription process. However, because only words were transcribed, I only used the square brackets to refer to the researcher's descriptions rather than transcriptions; for example, I used the following phrase whenever the participants interrupted me: [interrupted me]. Each interview was fully transcribed and then subsequently translated into English. Despite the advantage of being an Arabic native and from the same city as the participants (thus grasping the cross-cultural meanings and being able to cope with the Arabic dialect of the participants), I went through the English versions and checked their accuracy against the recordings. Each participant was given a different name to keep her anonymous.

3.4. Data collection and analysis

The current study investigated the participants' attitudes towards the use of corpus in the classroom, so reflective forms and interviews were used to elicit this information. The data were analysed quantitatively and qualitatively; for example, the closed-ended questions in the reflective form were coded as scores, while the open-ended questions were coded according to the themes that emerged from the data. The data obtained from the interviews were analysed qualitatively by drawing out themes and mapping the participants' responses. Dörnyei (2007) described two approaches for analysing interviews: the use of specific methodology (grounded theory) and the use of a more general methodology (the generic approach). Most published research has favoured the generic approach, which is a 'qualitative content analysis' (Dörnyei, 2007, p. 245). Therefore, I applied the generic approach to analyse the data because the grounded theory is

always used to generate or delimit a theory, and the interviews in the current study were constructed for exploratory purposes.

Dörnyei (2007) mentioned four phases in the analytical process of the generic approach: transcription, precoding and coding, growing ideas or memos, and interpreting the data. The transcription phase includes transcribing the data by listening to the recordings and typing all the interviews. The second phase – precoding and coding – consists of three levels: initial coding, second-level coding and using a template of codes. I excluded this third level because it requires a predetermined code template, which contradicts the more flexible nature of the interviews. Therefore, the first level was the initial coding of the data, and then, the codes were clustered into categories and finally factors. The third phase – growing ideas or memos – was performed at the second level during hard copy transcription. The fourth phase, interpreting the data by discussing the results, which can be found in the discussion section. The analysis used in the present study emphasises common factors and themes by producing frequency scores of the given answers. A necessary step in the analysis of qualitative data is to thematically organise answers prior to summarising the data; this step is important if useful conclusions are to be drawn from qualitative data (Moore, 2000; Cohen et al., 2011). Once each type of data was coded, I inputted the coded data into an Excel spreadsheet and then imported it into IBM SPSS (22.0) for analysis.

4. Results

4.1. The ability to use a computer and its effect on the use of the DDL approach in the classroom

The first research question aimed to ascertain if there was a relationship between the participants' computer skills and their performance while using the DDL approach; it also set out to determine if the participants' computer skills affected their use of the DDL approach, as well as their attitudes towards the DDL approach. Therefore, the quantitative data (reflective form) and qualitative data (open-ended questions from the reflective form and interviews) were used to address this research question.

First, it was important to know how the participants rated their own abilities with computers. As revealed in the reflective form, 49% of the participants described themselves as very competent computer users, 45% said they were quite competent users, and 6% said they were not very competent. This can be explained by the fact that 98% of the participants had their own computers at home and were quite capable of using them.

Second, to identify the relationship between the ability to use a computer and the use of the DDL approach, I ran a correlation analysis between two variables: the ability to use computers and the time spent on the sessions. To this end, I conducted a Spearman's rho test because the data were non-normally distributed. There was a small positive correlation between the ability to use computers and the time spent on the sessions, $r = .28$, $n = 51$, $p < .05$. It is worth noting that the performance of the three participants who described themselves as noncompetent computer users (participants 18, 47, 51), did not exhibit lower scores than the other participants who described themselves as very competent users in utilizing the steps of investigating the concordance lines. For example, participant 18 outperformed in the process performance (100%), with participants 47 and 51 also performing well in the process performance (92% and 88.5%, respectively). Some participants (e.g., 7 and 40) described themselves as very competent computer users and had 80.7% and 90% successful completion rates in the tasks, respectively. Additionally, the participants (e.g., 27 and 19) who described themselves as quite competent users were at 84.6% and 88% successful completion rates, respectively. Therefore, the ability to use a computer showed no relation to the process of investigating concordance lines.

Also, in the interviews (where the participants were given the anonymised names Amel, Hind, Lyla, Soha, Eman and Salam for ease of reference), all the participants had declared that their ability to use a computer had no correlation to either their performance or their attitude towards the DDL approach. Lyla and Eman explained that their attitudes and abilities were related to the training they received rather than their ability to use computers. As Lyla observed, *'I am not a good computer user; my*

ability is very poor, but I love using the corpus to learn new words. The ability to use the computer had no effect as I am a very poor user, and I could manage to use the corpus'. Eman reinforced this stance by saying, 'I am not a good user. No, it is not related to the use of corpus. The software is clear and easy, you follow clear instructions, so you do not need to be a very competent user of computers'.

4.2. Participants' overall evaluations of the use of the DDL approach and the corpus resource in the classroom

The second research question aimed at investigating the participants' evaluations and attitudes towards the use of the DDL approach by analysing the quantitative data (the reflective form) and qualitative data (open-ended questions from the reflective form and the interviews). The analysed data were divided into the following main themes: the use of the DDL approach, the general view of the corpus resource (the concordancer) and final thoughts. For each theme, the analysis encompassed the quantitative data (reflective form) and qualitative data (open-ended questions from the reflective form and interviews).

The use of the DDL approach

In the reflective form, 53% of the participants believed that the use of the DDL approach in the tasks was very easy, 23.5% thought it was quite easy and 23.5% thought it was easy. This point was further discussed in the interviews, with all six participants agreeing on the ease of the tasks when using the DDL approach. Regarding how helpful the approach was in reading and analysing the concordance lines, 57% of the participants strongly agreed, 33% agreed and 6% agreed to some extent that it was helpful. The remaining 4% did not think the approach was helpful in reading and analysing the concordance lines. Given this feedback, it was necessary to investigate in detail why they had arrived at these evaluations. The participants were asked what advantages and disadvantages they experienced in using the DDL approach.

Based on frequency scores, their answers were analysed to determine common factors and themes. I first identified the key themes and then counted their ratio frequencies, considering all the emerging themes (even if only mentioned once). These themes were divided according to their number of co-occurrences by separated participants, as shown in Table 1.

Table 1. The Emerged Themes from the Qualitative Data (Reflective Form)

	Themes	Co-occurrences	Percentage	Total No.
Advantages	New method for learning new words/meanings	Repeated twice	3.9%	51
	Help increase/improve the knowledge of words and meanings of words	Repeated 26 times	50.9%	51
	Smooth and easy use	Repeated 16 times	31.37%	51
	Helps to understand the meanings of words according to their collocates	Repeated 17 times	33%	51
	Knowing new words with no meanings	Repeated 6 times	11.7%	51
	Facilitate searching for new meanings of a word	Repeated 7 times	13.7%	51
	General theme of usefulness	Repeated twice	3.9%	51
	Improve English language	Repeated 14 times	27.45%	51
	Help in building sentences	Repeated twice	3.9%	51
Disadvant	No translation feature	Repeated	3.9%	51

ages	twice		
No pictures	Repeated 4 times	7.8%	51
The sentences being too long	Repeated 6 times	11.7%	51
Time-consuming	Repeated 9 times	17.6%	51
The need for training and practice	Repeated 7 times	13.7%	51
Some contexts were difficult because they did not have enough knowledge of the words in that context	Repeated 12 times	23.5%	51
No disadvantages of the approach	Repeated 19 times	37.2%	51

The disadvantages emerging from this feedback predominantly focus on more technical issues, though some themes conveyed a general criticism about using the DDL approach in the classroom, such as the sentences being too long, the approach being time-consuming and a need for training and practice. Another point was related to the participants' level, where it was discussed that some contexts were difficult because they did not have enough knowledge of the words in the specific context.

In the interviews, all six participants found that the DDL approach was easy and direct. Soha emphasised that it was easy after the first training session '*because all sessions were repeated but with different verbs*'. Hind explained this ease of use being a result of the directness of the approach: '*It was a step-by-step approach that led me directly to the answers*'.

The six participants expanded on their responses from the questionnaire by explaining why they found this approach to be helpful. For example, Salma said that the approach was helpful because of its gradualness in dealing with the concordance lines in

a *'step-by-step approach'*. Eman said that it helped her understand how the meaning of a noun affected the meaning of a verb. Lyla emphasised the approach's usefulness by drawing attention to the word on the right of the verb as a way for noticing the meaning of the verb and noun instead of the verb alone: *'It helped a lot, especially the sort feature. It was excellent. Using colours to notice the words to the right of the verbs was very good, as distinguishing meanings was easier'*. Hind described the approach as a *'neat approach'* that assisted her in figuring out the answers. Finally, Amel thought this approach supported learning actively by giving them an active role in the learning process:

It was a very engaging approach that made us think of the words by connecting them to other words. While using this approach, I felt that I was very active, and I really benefited from the sessions. ... We are not passive learners; we think while we are learning.

To gain a deeper understanding of the types of difficulties experienced by the learners, the third interview question was formulated to ask about the difficulties they experienced in the sessions. Five (out of the six) participants experienced technical difficulties in downloading the data file. This difficulty appeared in the training sessions only as they coped with it in the testing sessions. Another difficulty was related to the use of the DDL approach in answering the tasks that appeared in the testing sessions. Amel and Salma reflected that they needed time to comprehend some lines, which meant that some of the lines were difficult: *'Sometimes, we took time to comprehend the meaning of the target item. Some lines helped directly to figure the meaning, while others were tricky, such as the lines get revenge'* (Amel). Hind stated, *'Doing the tasks and using corpus without help was a bit challenging, especially in the first testing session. Then, it was easy'*. This statement could possibly be interpreted as initial difficulties in taking control of the concordances. Lyla also found difficulty in some lines: *'My [proficiency] level is not very good [low], so my vocabulary knowledge is limited. Therefore, I found some lines to be difficult or tricky. Other than that, everything was good and clear'*. Soha attributed her difficulties to her low-intermediate level: *'I did not know enough words in the context. If my level was higher, I*

would not have had this difficulty'. Also, she thought that the final activity was difficult because of the type congruency of the targeted collocations.

Given the general acknowledgement that the DDL approach supports independent learning (e.g., Bernardini, 2004), the participants were asked if they preferred guidance throughout the sessions or if they would have opted to use the corpus independently (the fourth question). All six participants concurred that after receiving the training, they were able to use the corpus and answer the activities themselves independently. Amel, Lyla and Salma said that after the two sessions of training, the tasks were manageable and not difficult when using the corpus, *'but if we had tried to do it for the first time by ourselves, it would have been difficult. We did not need help or guidance after we had the training'* (Amel). Salma observed, *'It will be good to provide help upon request'*. However, Hind, Soha and Eman said they were able to use the corpus independently but had trouble in the translation activity because they sometimes did not know the exact translation of the collocation. As Hind explained, *'Expressing the meaning of the word was a bit difficult'*. Soha explained this difficulty in greater detail: *'We knew that the verb has an exact meaning, but when we read the lines, it did not make sense. When the verbs are followed by certain nouns, the meaning changes'*. The participants related this difficulty to the exercise itself, but I believe it may have been related to the nature of the target collocation because the target collocations were incongruent, and the participants could not find an exact translation (e.g., *do homage*). The participants were additionally asked if they preferred help and guidance in the testing sessions; most of them preferred independent learning and declared that after the training sessions, there was no need for help or guidance from the instructor because the training was sufficient. Amel said, *'It would be nice to have help, or at least to know that we could get help if we asked for it'*. However, Lyla noted, *'After receiving the training, there is no need for help. Some contexts were still difficult, but this was related to my level, so the guidance would not have helped'*.

General view of the corpus resource (AntConc, concordancer)

Regarding the use of the corpus tool in the future, the participants were divided. From the 51 participants surveyed, 35% were very sure they would use the corpus tool, 39% would probably use it in the future, 22% could use it in the future (possibly) and 4% declared they would probably not use it. Most of the participants who said they would use the corpus tool in the future provided the following reasons for doing so: the tool is very useful for knowing new meanings of words (repeated 11 times), the tool helped in learning how to use words correctly and using words in the right context (repeated six times), and the tool was useful in developing English language (repeated eight times). Other reasons were rather general, such as the tool was easy to use (repeated six times), it was very useful (repeated eight times), it was enjoyable (repeated six times), the tool was good (repeated once), and they loved it (repeated five times); their generality of the responses, while vaguely positive, does not indicate the recognised pedagogical benefits of the approach.

Some of the participants who said that they would possibly use the corpus tools mentioned negative reasons for this possibility, such as using it only if needed [*have to use it*] (repeated eight times), too time-consuming to use (repeated three times), boring (repeated once) and requiring a lot of training and practice (repeated once). Other participants that said they could use it mentioned reasons such as boring with limited usefulness (repeated seven times) and a lack of time to use it in (repeated once). Seven participants did not give any reasons.

The participants' responses in the interviews were similar to the questionnaire responses, with five participants showing a desire to use it in the future. Their responses featured similar reasons (from the questionnaire), such as '*It was very effective to learn the words from the context*' (Amel). Hind justified her continued interest based on her current needs: '*I will also use it for learning English that is unrelated to curriculum, for example, how to write letters. I will use it according to my needs*'. This view was balanced by Salma, who, although admiring the concept of the corpus and the

approach followed in the sessions, said she would not use it in the future because she did not like the AntConc software: *'The software interface seems old with no lively colours.... I hate the idea of downloading the data file.... It does not support pictures or video clips'*. Salma did not object to trying other software given that the DDL approach is a new method for learning English within a native context.

The attitudes towards the concordancer were also explored in the interviews. All six participants agreed on the great benefits of using the concordancer and how using it was a good experience; they also thought that using the concordance lines to find answers for the activities and that finding different meanings of the target items were beneficial. As Amel explained, *'It helped in answering the tasks by connecting the meaning of the verb to the nouns... and made me think deeply about the target words, one word can have more than one meaning'*. However, Hind said that using the concordance lines *'was a little bit difficult. Some lines were tricky, but I managed to do well, other lines were not difficult'*. She further attributed this trickiness to the fact that it was a new method for her. Salma similarly found that some lines were not helpful in learning the meanings of the target words, reasoning that she would not have this difficulty if her level was higher.

Ultimately, all six participants believed that reading the concordance lines prior to the activities was a very good and useful way to draw attention to the repeated patterns, as well as to the different meanings a verb can have according to its collocates. As Amel said, *'It drew my attention to the repetition and the forms as well as the differences in meaning, which was better than giving us the information'*, while Hind observed that *'it was a good way to notice things [patterns] by myself'*. However, Eman stated, *'It was a good way, but sometimes, it was boring, so I read only some of the lines. Once I got the meaning, I stopped reading the lines'*. While Soha said, *'It was useful to draw our attention to the repeated patterns. But I did not read the lines first; I looked at the tasks, then read the lines'*.

Finally, regarding the closing question from the interview, where the participants were allowed to further expand on their answers, all the participants believed it was a great opportunity for learning English. Furthermore, many comments were made about the benefits they could gain by using the corpus for English specific purposes (ESP), specifically for learning medical terminology and understanding English used in the medical field. Amel suggested designing an English curriculum and including the use of the corpus as an activity session. Two participants also suggested developing the corpus tool as an application that could be used on their phones. Another participant thought that downloading the data file was tricky, so a file-sharing feature would be a good idea for further developing the software. Finally, one participant believed that using this experiment for higher levels could be even more promising because there would not be as much difficulty in understanding the context of the target items.

5. Discussion

5.1. Computer skills and attitudes towards corpus use in the classroom

As shown in the findings, the ability to use a computer had no effect on the learners' attitudes or their ability to use the DDL approach. Interestingly, computer skills may have had an effect on the time spent using the corpus resource because a small positive correlation was evidenced between the ability to use computers and the time spent on the sessions. This could mean that the participants would spend less time completing the tasks as they became more confident in their use of computers and would hesitate less while performing the steps.

Boulton (2010a, p. 534) argued that 'taking the computer out of the equation' can eliminate potential barriers to using the DDL approach for novice learners. He also recommended using paper-based materials instead of hands-on concordancing, specifically for low-level learners. Other researchers have identified technological difficulties as a source of the problem in using DDL in the classroom (e.g., Yoon & Hirvela, 2004). In contrast, Ito et al. (2009) mentioned how this generation spent their youth in direct

contact with the media, so media use is a main part of their lives. Pérez-Paredes et al. (2011, p. 246) noted that because the ‘guided-corpus consultation does not detract hands-on, activity time’, the use of digital media and online communication continues to increase. In the present study, this finding was also confirmed by the interviews with the six participants, who agreed that their attitudes and their ability to smoothly use the DDL approach were related to the training they had received, not to their ability to use computers. From the interview data, two participants admitted to being poor users, but they liked using the corpus (AntConc), thus supporting the argument that computer skills have little effect on the use of corpora in the classroom. The more important aspect here, however, is to provide guidance and training to gain corpus literacy, as Götz and Mukherjee (2006) recommended.

5.2. The participants’ overall evaluations of the use of the DDL approach/corpus recourse in the language classroom

The second question dealt with the learners’ evaluations. The overall evaluations of the DDL approach (from the questionnaire data) were highly positive, with all the participants agreeing that the DDL approach was easy to use. The participants’ positive attitudes towards the approach and the use of concordance lines for learning GVs were confirmed through the interviews. All six of the interviewees found that the DDL approach was straightforward and easy to use. These findings are consistent with the results reported in other studies, such as the results by Varley (2009), Yoon and Hirvela (2004), Chambers (2005) and O’Sullivan and Chambers (2006).

In the present study, the participants agreed that the approach was helpful in reading and analysing the concordance lines; this is consistent with Varley’s (2009) study, whose findings show that learners had positive responses to the corpus consultations and that they were able to identify the benefits of those consultations, particularly in terms of vocabulary acquisition. From the follow-up interviews in the present study, the participants agreed on the benefits and value of using the concordancer; they thought that

using the concordance lines was helpful in finding answers for the activities and in determining the different meanings of the target items. These positive attitudes were extended to the participants' desires to use the DDL approach in the future although they were split on the degree of how confident they were about using it, which was consistent with previous studies (e.g., Varley, 2009; Yoon & Hirvela, 2004).

Numerous studies have noted the importance of training (e.g., Chambers, 2005; Cheng et al., 2003; O'Sullivan & Chambers, 2006; Yoon & Hirvela, 2004), and the present study's interviewees also declared that the training was effective and efficient. All of the interviewees explained that after receiving the training, they were able to use the corpus and answer the activities independently. Moreover, after the training, they found the tasks to be manageable and not difficult when using the corpus.

The participants were also able to identify some of the advantages of using the DDL approach, which supports the importance of using concordance lines with guided training for learning GV patterns; doing this may help learners improve their knowledge of the meanings of words, understand the meanings of words according to their collocates and, in general, help improve their use of the English language.

In fact, most of the participants did not mention any disadvantages related to using the DDL approach. The few disadvantages that were mentioned were consistent with the disadvantages noted in previous research, predominantly noting that the contexts (concordance lines) were difficult because the participants did not have enough knowledge of the words in the specific context and that the sentences were occasionally too long. Some of the participants described the difficulty of the lines as '*the trickiness of lines*', in that some of the lines were not helpful in determining the meaning of the target items. They identified their proficiency level as a reason for their difficulty in being able to use the lines.

This disadvantage was also reported as a type of difficulty by Yoon and Hirvela (2004) and was more often mentioned by participants in the intermediate group than in the advanced

group. Another disadvantage relates to the nature of using corpus tools in a classroom because these tools are time-consuming and require training and practice (cf. O'Sullivan & Chambers, 2006). Technical difficulties were also reported in the interviews; especially, downloading the data file. This difficulty, however, was only associated with the training sessions because the participants were able to handle this task in the testing sessions.

A different type of difficulty was related to answering the tasks in the testing sessions, and this was subsequently mentioned in the reflective forms. This impacted the participants' evaluation of a task – here during the translation activity – in which they faced difficulty because they sometimes did not know the exact translation of the collocation. The participants related this difficulty to the exercise itself, but I believe it may be related to the nature of the target collocation because the target collocations were incongruent, and the participants could not find an exact translation. As expected, incongruent collocations were difficult for intermediate-level learners (the context of the present study), with GV patterns being problematic for L2 learners in general (Boers et al., 2014). For example, Laufer and Waldman (2011) found no differences in productive knowledge between lower- and upper-intermediate groups, with the most typical errors being the misuse of the verbs in those patterns, such as *do a mistake* instead of *make a mistake*. This issue is related to the L1 effect when the nouns collocate with different verbs; Nesselhauf (2005) and Yamashita and Jiang (2010) discussed this extensively.

In summary, the present study is significant in that it provided learners with effective guidance and direct steps to avoid difficulties concerning a lack of knowledge and skills in using corpora and concordancers – an issue that many previous studies failed to address. As well as the difficulties concerned with working on corpus data (Cheng et al., 2003; Lavid, 2007), and to avoid the difficulty they have faced, in general, when using the corpus for the first time.

6. Conclusion

In conclusion, the results of the students evaluations from the current study were quite positive. The interviews indicated that the six students – who represented various levels of proficiency – were able to make generalisations from the concordance lines. The present research has shown that using a concordancer can be successfully implemented in the Saudi context although some modifications are needed according to the students' suggestions. For example, integrating the use of a corpus in the English curriculum as an activity session could be implemented. The findings have shown that classroom concordancing can draw the student's attention to the notion of collocations and patterns, which can increase learner's proficiency. Inevitably, there are limitations in this research design. The present study is specific to the context of Saudi Arabian higher education; thus, it is important to be careful when interpreting the results. A typical limitation of all classroom research, is that the results need to be interpreted with caution due to the low number of participants. The study sample was 51 participants, which limits the generalisability of the results. Therefore, partial replications with larger-scale, different proficiency levels and target items are needed to increase the generalisability of the results.

Reference:

- Bernardini, S. (2004). Corpora in the classroom: an overview and some reflections on future developments. In: J. Sinclair (Ed.), *How to use corpora in language teaching* (pp.15-38). Amsterdam: John Benjamins Publishing company
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb–noun collocations. *Language Teaching Research*, 18(1), 54-74.
- Boulton, A., & Pérez-Paredes, P. (2014). Researching new uses of corpora for language teaching and learning. *Editorial to ReCALL*, 26(2), 121-127. DOI: 10.1017/S0958344014000068.
- Boulton, A. (2010a). Data-driven learning: on paper, in practice. In: T. Harris & M. Moreno Jaen (Eds.), *Corpus linguistics in language teaching* (pp.17-52). Bern: Peter Lang.
- Boulton, A. (2010b). Language awareness and medium-term benefits of corpus consultation. In A. Gimeno Sanz (Ed.), *New Trends in Computer-Assisted Language Learning: Working Together* (pp. 39-46). Madrid: Macmillan ELT.
- Boulton, A. (2010c). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534-572.
- Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153–172.
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language learning & technology*, 9(2), 111-125.
- Chambers, A., & O’Sullivan, I. (2004). Corpus consultation and advanced learners’ writing skills in French. *ReCALL*, 16 (01), 158-172.

Chan, T. P., & Liou, H. C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231-251.

Cheng, W. (2012). What can a corpus tell us about language teaching?. In: A. O'keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 319-332). Oxon: Routledge.

Cheng, W., Warren, M., & Xun-feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31, 173-186.

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing?. *System*, 25(3), 301-315.

Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 478-797). Cambridge, UK: Cambridge University Press.

Cohen, L, Manion, L., & Morrison, K. (2011). *Research methods in education* (5th ed.). New York: Routledge.

Cresswell, A. (2007). Getting to know connectors? Evaluating data-driven learning in a writing skills course. *Language and Computers*, 61(1), 267-287.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25-43.

Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching*, 45(4), 475-489.

Fox, G. (1998) Using corpus data in classroom. In: Tomlinson, B. (Ed.) *Materials development in language teaching* (pp. 25-43). Cambridge: Cambridge University Press.

Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam: Benjamins.

Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26(02), 225-242.

Götz, S., & Mukherjee, J. (2006). Evaluation of data-driven learning in university teaching: A project report. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods* (pp. 49-67). Frankfurt: Peter Lang.

Granath, S. (2009). Who benefits from learning how to use corpora?. In: K. Aijmer (Ed.), *Corpora and language teaching* (pp. 33-47) Amsterdam: John Benjamins Publishing Company.

Hafner, C. A., & Candlin, C. N. (2007). Corpus tools as an affordance to learning in professional legal education. *Journal of English for Academic Purposes*, 6(4), 303- 318.

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

Ito, M., Baumer, S., Bittanti, M., Boyd, D., Cody, R., Herr-Stephenson, B., Horst, H., Lange, P., Mahendran, D., Martínez, K., Pascoe, C., Dan Perkel, D., Robinson, L., Sims, C., & Tripp, L. (Eds.). (2009). *Hanging out, messing around, and geeking out: Living and learning with new media*. Cambridge, MA: The MIT Press.

Johns, T., & King, P. (1991). Classroom Concordancing. *English Language Research Journal*, 4, 1-16.

Johns, T. (1986). Microconcord: A language-learner's research tool. *System*, 14(2), 151-162.

Johns, T. (1991a). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *English Language Research Journal*, 4, 27-45.

Johns, T. (1991b). Should you be persuaded: Two examples of data-driven learning. *ELR Journal*, 4, 1-16.

- Johns, T. (2002). Data-Driven Learning: The Perpetual Challenge. In B. Kettemann & G. Marko, (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 107-117). Amsterdam: Rodopi.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language learning & technology*, 5(3), 77-90.
- Laufer, B., & Waldman, T. (2011). Verb-Noun Collocations in Second Language Writing: A Corpus Analysis of Learners' English. *Language Learning*, 61(2), 647- 672.
- Lavid, J. (2007). Contrastive patterns of mental transitivity in English and Spanish: A studentcentred corpus-based study. In E. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 237-252). Amsterdam: Rodopi.
- Leńko-Szymańska, A., & Boulton, A. (2015). Data-driven learning in language pedagogy. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple Affordances of Language Corpora for Data-driven Learning* (pp. 1-14). Amsterdam: John Benjamins Publishing Company.
- Moore, N. (2000). *How to do research: The Complete Guide to Designing and Managing research projects* (3rd ed.). London: Library Association Publishing.
- Nation, I. S. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge university press.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company.
- O'Sullivan, Í., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of second language writing*, 15(1), 49-68.
- Pérez-Paredes, P., Sánchez-Tornel, M., Alcaraz Calero, J. M., & Jiménez, P. A. (2011). Tracking learners' actual uses of corpora: guided vs non-guided corpus consultation. *Computer Assisted Language Learning*, 24(3), 233-253.

- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge: Cambridge university press.
- Richards, K. (2003). *Qualitative inquiry in TESOL*. UK: Palgrave Macmillan.
- Sinclair, J. (1999). A way with common words. In H. Hasselgard & S. Oksefjell (Eds.), *out of corpora: studies in Honour of Stig Johansson* (pp. 157-179). Amsterdam: Rodopi.
- Sinclair, J. (2003). *Reading Concordances: An Introduction*. London: Longman.
- Sinclair, J. (Ed.). (1987). *Looking Up: An Account of the Cobuild Project in Lexical Computing*. London: Collins.
- Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL*, 26(2): 184–201.
- Sun, Y-C., & Wang, L-Y. (2003). Concordancers in the EFL classroom: cognitive approaches and collocation difficulty. *CALL*, 16(1), 83-94.
- Varley, S. (2009). I'll just look that up in the concordancer: integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, 22(2), 133-152.
- Vyatkina, N. (2016a). Data-driven learning for beginners: The case of German verb- preposition collocations. *ReCALL*, 28(2), 207-226.
- Vyatkina, N. (2016b). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning and Technology*, 20(3), 159-179.
- Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44(4), 647-668.

Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning & Technology*, 12, 31-49. Retrieved from <http://llt.msu.edu/vol12num2/yon.pdf>.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of second language writing*, 13(4), 257-283.