Corpora and Concordancing in Second Language Vocabulary Learning

Morad Alsahafi

Department of European Languages & Literature, Faculty of Arts and Humanities, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract—**Vocabulary knowledge is a core component of second language (L2) proficiency. The critical importance of vocabulary learning is acknowledged by L2 learners, teachers, and researchers alike. Vocabulary development is a growing area of interest in second language research in our increasingly technologized world where rapid technological advancement and widespread access to high-speed internet have started to provide unique and promising opportunities for language teaching and learning. This paper aims to present an overview of computer corpora and concordancing and their potential role in second language vocabulary learning. After providing a brief introduction to the construct of vocabulary knowledge and the role of computer technology in language learning, the paper focuses on the use of computer corpora and concordancing in L2 vocabulary teaching and learning. The paper then presents an example of a KWIC (Keyword in Context) concordance display in order to demonstrate the kinds of lexical information that a number of concordance lines can provide about the target word. Some concordance-based vocabulary activities are also provided. This paper is hoped to help L2 researchers, teachers and learners become more aware of the potential role of computer corpora and concordances in the development of L2 vocabulary knowledge in our highly technologized world.**

Index Terms—**corpus, concordance, KWIC display, vocabulary knowledge, vocabulary teaching and learning**

I. INTRODUCTION

Second language (L2) vocabulary knowledge represents a core component of L2 proficiency (McCarthy, 1990; Nation, 2013; Schmitt, 2000; Webb & Nation, 2017). The critical importance of vocabulary, as the building blocks of language, is acknowledged by L2 learners, teachers, and researchers alike. As Schmitt has rightly pointed out "one thing that students, teachers, materials writers, and researchers can all agree upon is that learning vocabulary is an essential part of mastering a second language" (2008, p. 329). Alderson (2000) notes that vocabulary knowledge is the "single best predictor of text comprehension" (p. 35).

Vocabulary learning is a growing area of interest in second language research. It is widely recognized in the field of vocabulary acquisition that vocabulary knowledge is a multi-dimensional construct incorporating several aspects of lexical knowledge related to the word's form (spoken, written, word parts), meaning (form and meaning, concept and referents, associations), and use (grammatical functions, collocations, constraints on use) (Nation, 2013). Research has suggested that L2 vocabulary can be learned either incidentally (implicitly) or intentionally (deliberately) (Lindstromberg, 2020; Nation, 2013; Webb, 2020). Intentional vocabulary learning is "the acquisition through activities designed to commit components of vocabulary knowledge to memory" whereas incidental learning is "the by-product of activities, usually involving comprehension, which are not explicitly designed for vocabulary learning" (Rodgers, 2018, p. 195).

According to Nation (2013), there are four kinds of vocabulary that may be encountered by L2 learners in any given text:

- High-frequency words: These include function words (e.g., prepositions, articles, etc.), and content words (e.g., nouns such as government, conference, production, etc.).
- Academic words: This includes common words in different kinds of academic texts (e.g., Coxhead's (2000) Academic Word List).
- Technical words: These are closely related to the topic and subject area of the text (e.g., a dictionary of economics, ICT, etc.).
- Low-frequency words: These are the biggest group of words. They include all the words that do not belong to the three groups listed above.

In recent years, rapid technological advancement and widespread access to broadband internet around the world have started to provide unique and promising opportunities for language teaching and learning. In this regard, the advent of computer assisted language learning (CALL) has offered a new dimension to the field of second language acquisition. Vocabulary learning and teaching are no exception. Examples of computer assisted learning used for L2 vocabulary development include the use of computer corpora and concordances, the development of specialized word lists through the use of corpus-based frequency counts, and the creation of on-line interactive vocabulary exercises. In this paper, the

focus is on the role of technology in L2 vocabulary learning. In particular, this paper aims to shed light on one main area in which technology can be used to enhance vocabulary learning both inside and outside the L2 classroom, namely the use of corpora and concordances. The following section provides a brief discussion of the role of technology in language teaching and learning.

II. THE ROLE OF TECHNOLOGY IN LANGUAGE TEACHING

What kind of a role does computer technology play in language teaching and learning? In recent years, computer technology has transformed the landscape of language learning and teaching (Levy, 2012; Richards, 2015). Students and teachers alike are expected to be technologically literate to take advantage of the opportunities that technology can provide for language development. Richards (2015) highlights the potential benefits that technology can provide for language learners and teachers, including providing a greater exposure to the target language and culture, creating a better learning environment which connects the classroom to the real world, and promoting active learning and learners' autonomy (see Table 1).

TABLE 1
TECHNOLOGY SUPPORT FOR LANGUAGE LEARNERS AND TEACHERS (RICHARDS, 2018, PP. 21-23)

Technology-supported learning	Technology-supported teaching
• *provides a greater exposure to the L2*	• *creates a better learning environment*
• *increases opportunities for authentic interaction*	• *enables more learner-centred teaching*
• *enables flexible learning*	• *supports teaching with mixed-level classes*
• *supports different skills and ways of learning*	• *expands the classroom to the real world*
• *suits learners of different proficiency levels*	• *enriches the curriculum*
• *encourages more active learning*	• *gives teachers a much wider range of strategies to use in teaching*
	• *facilitates student-learning monitoring*

The computer, in many cases, appears to play one of two major roles: tutor vs. tool (stimulus). Inside the language classroom, teachers can use the computer as a useful tool (a teaching aid) for creating language learning opportunities to promote the development of language skills (listening, speaking, reading, writing) and language components (vocabulary, grammar, phonology, graphology). Some of the common features of computer-based vocabulary learning programs are providing individualized feedback messages, providing automatic scoring, generating e-mail feedback messages, and the ability to randomize the order of questions to reduce answer sharing among students.

It is to be noted, however, that unless it is used by an experienced teacher both productively and imaginatively, "in and of itself technology is unlikely to make for good teaching" (Hampel & Baber, 2003, p. 189). Therefore, like in a lecture-based classroom, the role of the teacher remains to be crucial in CALL contexts. Different CALL programs entail different degrees of teacher intervention (Jones, 2001). According to Fox (1984, p.32), teachers are "unlikely to find the computer lightening their workloads, for it is a medium which offers new opportunities for innovative language teaching and learning." To ensure successful outcome of computer assisted language instruction, teachers should start by analyzing their own instructional goals as well as their students' needs and expectations. Additionally, teachers need to investigate what piece of technology can better address their teaching goals and their learners' and institutional needs. De Szendeffy (2005, p. 7) lists a number of CALL teachers' responsibilities, including:
- arrange project- or task-oriented activities
- manage the time spent on activities
- instruct students on effective use of applications
- guide students through resources for language learning
- encourage student-student interaction

In addition, the computer can be used as a tutor outside the language classroom. In this case, the computer plays a more complex role in the sense that it "instructs, provides content and exercises, guides the student, and gives feedback" (de Szendeffy, 2005, p. 10). For example, Groot (2000) describes a computer assisted vocabulary acquisition program (CAVOCA) in which the computer plays the role of the tutor by systematically taking the learner through a number of intentional vocabulary learning steps, i.e., noticing, storing the word in the memory by establishing lexical networks relationships with other words, and consolidation by means of further exposure to ensure long term retention. In sum, technology-enhanced language learning "provides limitless opportunities for…learning that draws on multiple modalities and that takes learners out of the classroom and into a world without walls" (Richards, 2015, p. 29). The next section discusses the use of computer corpora in L2 teaching and learning, specifically in vocabulary instruction.

III. COMPUTER CORPORA IN VOCABULARY LEARNING

Corpora or corpuses (singular corpus) can be defined as large electronic collections of spoken or written naturally occurring language (Ma & Mei, 2021; McEnery & Wilson, 2001; O'Keeffe et al., 2007; Reppen, 2010; Sinclair, 1991). Computer corpora "show how people use the language and provide objective evidence of fresh and authentic language use" (Ma & Mei, 2021, p. 179). Corpus research has been made possible thanks to the advancement in computer and software technologies, which gives the computer the ability to "store, code, categorize, and retrieve massive amounts of information" (Durand, 2018, p. 132). Furthermore, it goes without saying that computer technologies are able to provide more accurate and reliable corpus-based statistics than other traditional methods.

Based on corpus studies, a new area of linguistics, namely corpus linguistics, has been established (McEnery & Wilson, 2001). Generally, corpus linguistics plays two main roles in the field of language teaching and learning: "as a means to explore actual patterns of language use and as a tool for developing materials for classroom language instruction" (Reppen & Simpson, 2002, p. 92). Clearly, corpus linguistics has started to have its powerful impact on vocabulary learning research. "It is no exaggeration to say that corpora, and the study of corpora, have revolutionized the study of language, and of the applications of language, over the last few decades" (Hunston, 2002, p. 1).

Vocabulary learning and corpus analysis are closely related to each other due to the data-driven nature of corpora (Lin & Lee, 2015). Consequently, the field of vocabulary learning and teaching has been continually affected by corpus linguistics. Research on L2 vocabulary acquisition has found that the "integration of corpora into vocabulary classrooms not only provides learners with faster searching tools and better quality of contexts that traditional dictionaries are not likely to achieve but also enhances their learning motivation" (Chen, 2004, p.5).

There are many types of English language corpora. Some of the most widely known ones include:

- The Brown Corpus (full name: Brown University Standard Corpus of Present-Day American English): the first modern computer-based general corpus containing around one million words drawn from a wide variety of American English sources.
- The British National Corpus (BNC): a general corpus containing 100 million words of modern British English.
- The American National Corpus (ANC): a general corpus containing 22 million words of American English.
- The International Corpus of English (ICE): a specialized type of corpora containing samples of both written and spoken English language from around the world (e.g., ICE-GB, ICE-NZ, ICE-NIG). Each sub-corpus of the ICE comprises (or will comprise) one million words.

Broadly speaking, corpus linguistics research (see, for example, Barlow, 1996; Kaltenbock & Mehlmauer-Larcher, 2005; Reppen & Simpson, 2002) has identified two general uses of corpora in language teaching (Figure 1).
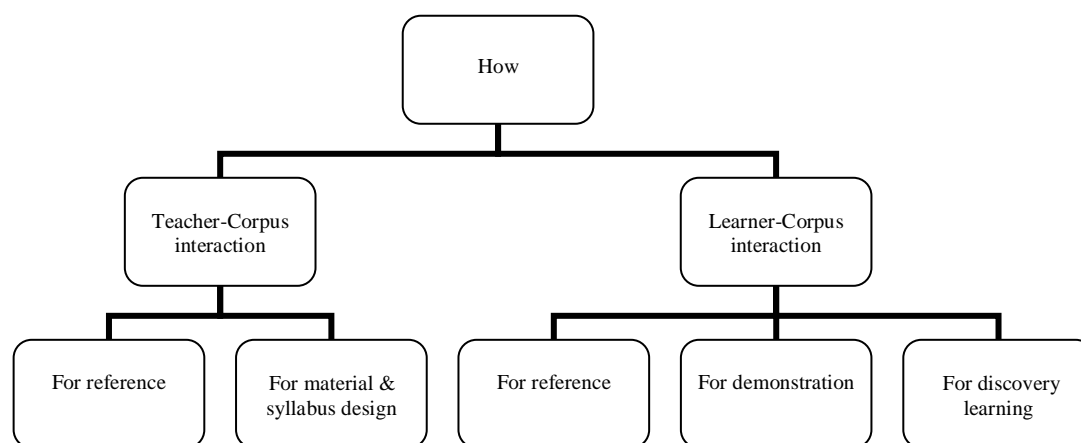

Figure 1 Uses of Computer Corpora in Language Teaching and Learning (Source: Kaltenbock & Mehlmauer-Larcher, 2005, p. 79)

As illustrated in Figure 1, there are two types of corpus-based interaction: teacher-corpus interaction and leaner-corpus interaction. The former is related to teachers' use of corpora (e.g., outside the classroom for reference purposes and syllabus design, or inside the classroom for reference or demonstration purposes). The latter is related to learners' use of corpus information. For example, learners can interact with corpora for inductive or data-driven kind of learning (e.g., for reference, demonstration, discovery, or exploratory learning purposes).

It should be noted that corpora can provide different levels of information about target words. Certain tools (e.g., MonoConc, WordSmith tools, Nation's Vocabulary programs) are needed to interact with the corpus. In general, we can identify two basic types of information that a corpus can provide:

- The first type is of information a corpus can provide is related to frequency of occurrence which represents one of the major contributions of language corpora to the language classroom (Kaltenbock & Mehlmauer-Larcher, 2005). For example, a list of the most frequent academic words can be generated based on frequency of occurrence (e.g., Coxhead's Academic Word List).

- The second type of information that a corpus can provide is based on the use of concordance tools which can be used to give lexical information about various uses and senses of the target word such as its lexico-grammatical patterns (Reppen & Simpson, 2002).

A number of researchers have discussed the advantages of using computer corpora as useful authentic resources in the vocabulary classroom (Boulton & Cobb, 2017; Daskalovska, 2015; Flowerdew, 2012; Gordani, 2013; Lee et al., 2018; Kaltenbock & Mehlmauer-Larcher, 2005; Ma & Mei, 2021; Pérez-Paredes, 2022; Sinha, 2021). Kaltenbock and Mehlmauer-Larcher identified the potential of corpora for language teaching and learning as follows:

- They can provide information not easily available from other sources.
- Language items retrieved from corpora are not isolated linguistic items but come together with their co-text.
- Corpora give access to a verity of different text types.
- Computer corpora and their search software offer unique search facilities.
- They can facilitate discovery learning and promote learner autonomy, provided their use is appropriately mediated by the teacher. (Kaltenbock & Mehlmauer-Larcher, 2005, p. 81).

Similarly, Boulton and Cobb (2017) indicated that corpus-based data-driven learning leads to "increased language sensitivity, noticing, induction, and ability to work with authentic data" (p. 349). Pérez-Paredes (2022) pointed out that the use of computer corpora enhances inductive language learning because it enables learners to broadly access rich and authentic language data. The following section deals with the use of concordancing for vocabulary teaching and learning. The section provides an example of a concordance display and presents a variety of concordance-based vocabulary learning activities.

IV. CONCORDANCE-BASED VOCABULARY LEARNING

Concordancing represents "the primary means of investigating language corpora" (Ballance, 2017; Rayson, 2015). A concordance is "a list of all of the occurrences of a particular search term in a corpus, presented within the context in which they occur – usually a few words to the left and right of the search term" (Baker et al., 2006, pp. 42-43). A concordance is a means of accessing a corpus in order to obtain a variety of lexical information on a certain target word such as its occurrence across different texts, and its grammatical, collocational, and semantic properties (Choksuansup & Tangpijaikul, 2017; Flowerdew, 2015; Kartal & Yangineksi, 2018; Reppen & Simpson, 2002). There are available a wide variety of useful and high-speed concordancing tools (e.g., MonoConc, WordSmith tools, MlultiConcord) that can be used to elicit information on counts of frequency of lexical items (Respect & Muzi, 2021). There are also available a number of useful websites that can be used for concordancing purposes (e.g., Compleat Lexical Tutor or Lextutor).

Concordance programs can be used to generate concordance lines for a particular word. A concordance line is a line of a text taken from a corpus displaying an occurrence of a certain target word (Lindquist & Levin, 2018). Such a display is known as key word in context or KWIC (O'Keeffe et al., 2007). In this regard, Hunston (1995) suggests two ways of presenting concordance lines to learners. Language teachers can either present all concordance lines for the target word or select a number of lines for students to explore and discover particular patterns and find common regularities.

For example, Figure 2 shows 12 selected concordance lines for the target word *benefit* based on Brown corpus. As can be seen from this concordance display, the search term (i.e., *benefit*) is at the center of each concordance line in the KWIC view making it easier to see which words tend to follow it and which words tend to come before it.

```
 1  In addition to the rich soil they  benefit  by feedings of manure water every
 2  s and leg-split lifters would both  benefit  enormously by practicing those var
 3  nk of pensions as an important job  benefit  factor anyhow and they're liable t
 4  ions, no paid holidays; they still  benefit  far less than the "other" 50 per c
 5   when the body is mature enough to  benefit  from it without danger of injury,
 6  e with the number of employees who  benefit  from your program?   Have you aud
 7  s the donor received some material  benefit  in return, such as tickets to a sh
 8   the emperor had to use it for the  benefit  of the living. The royal ritual ge
 9  eryday usable English should be of  benefit  to all who seek the truth.  B12 17
10  CORPS VOLUNTEERS are assured a tax  benefit  under the law creating the agency.
11  e Ekstrohm, I want to give you the  benefit  of every doubt. But you aren't exac
12   leave when it will be of personal  benefit  to her. Recurrent problems may be
```
Figure 2 Concordance Display of KWIC for the Target Word: Benefit

By observing this KWIC display of *benefit*, learners can gain a wide range of lexical information including:
- parts of speech: *benefit* can occur as:
  - a verb (lines 1, 2, 4, 5, 6)
  - a noun (line 3, 7, 8, 9, 10, 11, 12)
- verb + preposition:
  - *benefit* by (line 1)
  - *benefit* from (line 5, 6)

- noun + preposition:
    - *benefit* of (line 8, 11)
    - *benefit* to (line 9, 12)
- Collocations:
    - *benefit* enormously (line 2)
    - job *benefit* (line 3)
    - *benefit* factor (line 3)
    - material *benefit* (line 7)
    - tax *benefit* (line 10)
    - personal *benefit* (line 12)
- Idiomatic expressions:
    - for one's *benefit* (line 8)
    - give someone the *benefit* of the doubt (line 11)

Based on these concordance lines, a number of useful vocabulary learning activities can be designed in order to provide both receptive and productive learning opportunities to promote the learners' knowledge of the target word (see the Appendix for some suggested concordance-based vocabulary learning activities).

Obviously, such a small KWIC display of 12 concordance lines does not show all of the occurrences of the word *benefit* in the corpus. These concordance lines were purposely selected to show certain lexical properties of the target word. As illustrated above, concordance lines provide important information pertinent to the target word through the presentation of many samples of real and authentic pieces of discourse containing the target word. With appropriate teacher guidance, L2 learners can explore these concordance lines in order to improve the quality of their knowledge of the target word. In general, the learning process tends to follow a more inductive approach to vocabulary learning.

According to Pereira (2004), the use of concordancing offers students "a new approach to understanding several different aspects of the language they are studying, such as meaning and semantic disambiguation, real differences between near-synonyms, morphosyntactic aspects, and information on the uses of words in natural contexts and term in specialized discourse" (p. 118). Corpus-based vocabulary studies have focused mainly on the use of concordances and collocations when developing university-level L2 learners' writing skills (see Pérez-Paredes, 2022, for a recent review of research on corpora and data-driven language learning).

V. CONCLUDING REMARKS

This paper has provided an overview of the use of computer corpora and concordances in vocabulary teaching and learning. As argued earlier, corpus-based vocabulary learning activities represent a useful approach that can be used to foster the development of lexical knowledge. Mediation by the teacher should be regarded as "a necessary prerequisite for successful application of computer corpora in language teaching" (Kaltenbock & Mehlmauer-Larcher, 2005, p. 81). In this regard, Barlow (2004) notes that it is important for both teachers and learners to be aware of the relationship between corpus access and corpus analysis. That is, any decision made in the retrieval process of a corpus of text would inevitably give voice to some patterns in that text and at the same time exclude some others. To overcome such a problem, Barlow (2004, p. 206) emphasizes the importance of "considering a range of different ways of analyzing a text." Likewise, Ma and Mei (2021) state that "teachers should be very selective in dealing with concordance lines and be strategic in integrating corpus resources into classroom teaching to facilitate student language learning" (p. 189).

Finally, notwithstanding their potential usefulness, corpus resources and tools have not been widely used by language teachers and learners in the language classroom (Ballance, 2017; McCarthy, 2008; Pérez-Paredes, 2022). Balance (2017) discusses three main reasons for this seemingly limited uptake of concordancing in language learning contexts: lack of teacher training, limited access to technological tools and resources, and possible competitions between the cognitive demands involved in the process of language learning and the use of concordancing. According to McCarthy (2008), teachers' unfamiliarity with corpora and their lack of awareness on how to develop and design corpus-based activities have been the main reason for being neglected in the language classroom. Therefore, teachers training is needed to raise their awareness of the potential role of corpora in language learning as well as the integration of corpora into language syllabus (Pérez-Paredes, 2022).

As noted by Vyatkina and Boulton (2017, p. 2), research into corpora use in language learning is "still developing". In this connection, the use of corpora and concordances represents one of the most promising areas of research in computer assisted language learning (Ballance, 2017). Hence, more research is needed to enhance our understanding of the potential role of computer corpora and concordances in the development of L2 vocabulary knowledge. Meanwhile, this paper is hoped to help L2 researchers, teachers and learners become more aware of the potential role of computer corpora and concordances in the development of L2 vocabulary knowledge in our increasingly technologized world.

APPENDIX. CORPUS-BASED VOCABULARY ACTIVITIES

Look at the concordance lines provided below for the target word *benefit*. Pay special attention to what comes before and after the target word. Then answer the questions.

```
 1   In addition to the rich soil they benefit by feedings of manure water every
 2  s and leg-split lifters would both benefit enormously by practicing those var
 3  nk of pensions as an important job benefit factor anyhow and they're liable t
 4  ions, no paid holidays; they still benefit far less than the "other" 50 per c
 5   when the body is mature enough to benefit from it without danger of injury,
 6  e with the number of employees who benefit from your program?   Have you aud
 7  s the donor received some material benefit in return, such as tickets to a sh
 8   the emperor had to use it for the benefit of the living. The royal ritual ge
 9  eryday usable English should be of benefit to all who seek the truth.  B12 17
10  CORPS VOLUNTEERS are assured a tax benefit under the law creating the agency.
11  e Ekstrohm, I want to give you the benefit of every doubt. But you aren't exac
12   leave when it will be of personal benefit to her. Recurrent problems may be
```

1. Does the word "benefit" occur as a noun, a verb, an adjective, or an adverb?
    ……………………………………………………………………………………………
2. What prepositions often proceed "*benefit*"?
    ……………………………………………………………………………………………
3. What prepositions often follow "*benefit*"?
    ……………………………………………………………………………………………
4. What words (other than prepositions) does "*benefit*" often occur with?
    ……………………………………………………………………………………………
5. What does it mean to give someone "the *benefit* of the doubt"?
    ……………………………………………………………………………………………
6. Write one sentence using the following prepositional phrases.
    benefit from: ……………………………………………………………………………
    benefit by: ……………………………………………………………………………
    benefit to: ……………………………………………………………………………
    benefit of: ……………………………………………………………………………

REFERENCES

[1]   Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
[2]   Baker, P., Hardie, A., & McEnery T. (2006). *A glossary of corpus linguistics*. Edinburgh: Edinburgh University Press.
[3]   Ballance, O. J. (2017). Pedagogical models of concordance use: correlations between concordance user preferences, *Computer Assisted Language Learning*, *30*(3-4), 259-283.
[4]   Barlow, M. (1996). Corpora for theory and practice. *International Journal of Corpus Linguistics, 1*, 1–37.
[5]   Barlow, M. (2004). Software for corpus access and analysis. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 205-221). Amsterdam: John Benjamins.
[6]   Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348-393.
[7]   Chen, Y. H. (2004). The use of corpora in the vocabulary classroom. *The Internet TESL Journal, 10*(9), 1-5.
[8]   Choksuansup, N., & Tangpijaikul, M. (2017). The effect of concordancing-based teaching on vocabulary learning and retention on Thai EFL learners. *Chophayom Journal*, *28*(3), 11-19.
[9]   Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213-238.
[10] Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning, 28*(2), 130-144.
[11] De Szendeffy, J. (2005). *A practical guide to using computers in language teaching*. Ann Arbor, MI: University of Michigan Press.
[12] Durand, S. (2018). Embracing computer corpora in the language learning classroom and using it in your classroom. In J. Perren, K. Kelch, , J.Byun, S. Cervantes, & S. Safavi (Eds.), *Applications of CALL theory in ESL and EFL environments*. United States of America: IGI Global.
[13] Flowerdew, L. (2012). *Corpora and language education*. Basingstoke: Palgrave Macmillan.
[14] Flowerdew, L. (2015). Corpus-driven learning and language learning theories. In A. Lenko & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp.15–36). Amsterdam; Philadelphia, PA: John Benjamins.
[15] Fox, J. (1984). Computer-assisted vocabulary learning. *ELT Journal, 38*, 27-33.
[16] Gordani, Y. (2013). The effect of the integration of corpora in reading comprehension classrooms on English as a foreign language learners' vocabulary development. *Computer Assisted Language Learning, 26*(5), 430-445.
[17] Groot, P. J. M. (2000). Computer assisted second language vocabulary acquisition. *Language Learning & Technology, 4*, 60-81.
[18] Hampel, R., & Baber, E. (2003). Using internet-based audio-graphic and video conferencing for language teaching and learning. In U. Felix (Ed.), *Language learning online: Towards best practice* (pp. 171-191). Lisse: Swets & Zeitlinger.
[19] Hunston, S. (1995). Grammar in teacher education: the role of a corpus. *Language Awareness, 4*, 15-31.
[20] Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

[21] Jones, J. F. (2001). CALL and the responsibilities of teachers and administrators. *ELT Journal, 55*, 360-367.

[22] Kaltenbock, G., & Mehlmauer-Larcher, B. (2005). Computer corpora and the language classroom: on the potential and limitations of computer corpora in language teaching. *ReCall, 17*, 65-84.

[23] Kartal, G., & Yangineksi, G. (2018). The effects of using corpus tools on EFL student teachers' learning and production of verb-noun collocations. *PASAA: Journal of Language Teaching and Learning in Thailand*, *55*, 100-125.

[24] Lee, H., Warschauer, & M. Lee, J. H. (2018). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, *40*(5), 721-753.

[25] Levy, M. (2012). Technology in the classroom. In A. Burns & J. C. Richards (Eds.), *Pedagogy and practice in language teaching*. Cambridge: Cambridge University Press (pp. 279–86).

[26] Lin, M., & Lee, J. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal, 69*(3), 264 – 274.

[27] Lindquist, H., & Levin, M. (2018). *Corpus Linguistics and the Description of English* (2nd ed.). Edinburgh University Press.

[28] Lindstromberg, S. (2020). Intentional L2 vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 240–254). London: Routledge.

[29] Ma, Q., & Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *Journal of China Computer-Assisted Language Learning*, *1*(1), 177-190.

[30] McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.

[31] McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching, 41*(4), 563-574.

[32] McEnery, A. M., & Wilson, A. (2001). *Corpus linguistics: An introduction*. Edinburgh, UK: Edinburgh University Press.

[33] Nation, I. S. P. (2013). *Learning vocabulary in another language (2ⁿᵈ Ed.)*. Cambridge: Cambridge University Press.

[34] O'Keeffe, A., McCarthy, M. J., & Carter, R. A. (2007). *From corpus to classroom: language use and language teaching*. Cambridge: Cambridge University Press.

[35] Pereira, L. A. S. (2004). The use of concordancing in the teaching of Portuguese. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 109-122). Amsterdam: John Benjamins.

[36] Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning, 35*, (1-2), 36-61.

[37] Rayson, P. (2015). Computational tools and methods for corpus compilation and analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics*. Cambridge: Cambridge University Press.

[38] Reppen, R. (2010). *Using corpora in the language classroom*. New York: Cambridge language Press.

[39] Reppen, R., & Simpson, R. (2002). Corpus linguistics. In N. Schmitt (Ed.), *Introduction to Applied Linguistics* (pp. 92-111). London: Arnold.

[40] Respect, M., & Muzi, M. (2021). Using MonoConc Pro to teach and learn lexical collocations in Xitsonga. *Proceedings of the International Conference of the Digital Humanities Association of Southern Africa*, 1-7.

[41] Richards, J. C. (2015). Technology in Language Teaching Today. *Indonesia Journal of English Language Teaching, 10*(1), 18-32.

[42] Rodgers, M. (2018). The images in television programs and the potential for learning unknown words. Approaches to Learning, Testing, and Researching L2 Vocabulary. *ITL - International Journal of Applied Linguistics 169*(1), 191-211.

[43] Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

[44] Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329-363.

[45] Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

[46] Sinha, T. S. (2021). EFL learners' perception of and attitude to corpus as a vocabulary learning tool. *The Reading Matrix: An International Online Journal, 21*(2), 106-119.

[47] Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology, 21*(3), 1–8.

[48] Webb, S. & Nation, P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.

[49] Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). London: Routledge.

Morad Alsahafi is an Associate Professor of Applied Linguistics in the Department of European Languages and Literatures, King Abdulaziz University, Jeddah, Saudi Arabia. Dr. Alsahafi was awarded a PhD in Applied Linguistics from the University of Auckland, New Zealand in 2010. His research interests are in applied linguistics and sociolinguistics, particularly heritage language maintenance and education, bilingualism, language planning (in local contexts), and EFL teaching and learning. He has most recently published in International Journal of Bilingual Education and Bilingualism and Journal of Multilingual and Multicultural Development. His current research focuses on second language vocabulary learning and issues related to heritage/community languages in Saudi Arabia.