Global Scientific Review
A Peer Reviewed, Open Access, International Journal
www.scienticreview.com
Volume 48 February 2026 ISSN (E): 2795-4951

Designing A Web-Based Corpus Collocation Lab For Grade 8–9 Efl Learners: Principles, Structure And Classroom Use

Bahronov Nodirbek Aliyevich
Independent researcher, Department of Interfaculty Foreign Languages, Bukhara State University
e-mail: bahronovnodirbek@gmail.com

Abstract. This article describes the design of "CORPUS-BASED COLLOCATIONS' LAB", a website developed to support the corpus-based teaching of English collocations to grade 8–9 learners in general secondary schools of Uzbekistan. The paper outlines the pedagogical principles underlying the website, its structure based on the six-stage "Corpus Detective" exercise system (Notice – Search – Analyze – Practice – Produce – Reflect), the selection of collocations from the SkELL and COCA corpora, and the alignment of content with textbook units. Practical recommendations for using the website in classes with full, limited or no internet access are given.
Keywords: Web-based learning, corpus, collocation, website design, CALL, secondary school, EFL, Corpus Detective.

Introduction

Corpora offer teachers and learners rich, authentic evidence of how words combine. However, existing corpus interfaces are designed for researchers and university students rather than for 14–15-year-old learners. Boulton (2009) showed that lower-level learners can benefit from corpus work, but only when the materials suit their level. In practice, when school learners are sent directly to a large corpus, they often get lost in long lists of examples and unfamiliar terminology. There is therefore a need for an intermediate environment that preserves the benefits of corpus evidence while presenting it in a structured and age-appropriate way.

To meet this need, we developed "CORPUS-BASED COLLOCATIONS' LAB", a website that hosts collocation tasks for grades 8 and 9. The aim of this article is to describe the principles, structure and classroom use of the website.

Design principles

The design of the website was guided by principles drawn from computer-assisted language learning (CALL), multimedia learning and data-driven learning.

Learner fit. Chapelle (2001) lists learner fit among the key criteria for evaluating CALL tasks: the task must match learners' level and characteristics. Accordingly, corpus examples on the website are short, the interface uses simple English with Uzbek instructions where necessary, and each task focuses on a small number of collocations.

Reduced cognitive load. Following Mayer's (2009) principles of multimedia learning, each screen presents one task, irrelevant information is removed, and the node word is highlighted in concordance lines. This helps learners focus on the pattern rather than on the interface.

Guided discovery. In line with data-driven learning (Johns, 1991; Gilquin & Granger, 2010), learners are not given rules in advance. Instead, they answer

Global Scientific Review

questions about the examples and formulate the pattern themselves, after which the website shows a model answer.

Curriculum alignment. All tasks are built around textbook units so that corpus work reinforces the lesson rather than competing with it.

Structure of the website

The website is organised by textbook units. Grade 8 includes the units "Eating out", "Street food around the world", "The latest technology" and "Computers and the internet". Grade 9 includes "It's a challenge", "Our changing planet", "Cultural understanding: National parks", "Our holiday", "My place" and "Strange houses". Within each unit, tasks follow the six stages of the "Corpus Detective" exercise system (Table 1).

Table 1. Stages of the "Corpus Detective" system on the website

Stage	What learners do	Example (unit "The latest technology")
Notice	find highlighted collocations in a short text	*download an app*, *charge a phone*
Search	look up the key word in SkELL and record its collocates	*device*: *mobile device*, *use a device*
Analyze	group concordance lines and answer guiding questions	verbs vs. adjectives with *technology*
Practice	gap-fill, matching, "find the wrong collocation"	*I need to ___ my phone* (charge)
Produce	write a short review or record a spoken description	review of a new gadget
Reflect	check own collocations against the corpus; keep an error log	self-check list

Selection of collocations. For each unit, key nouns, verbs and adjectives were taken from the textbook. Their collocates were extracted with the "Word sketch" function of SkELL (Kilgarriff et al., 2015), and their frequency and genre distribution were checked in COCA (Davies, 2008–). Collocations were selected if they were frequent, relevant to the unit topic and useful for learners' own speaking and writing. Rare, idiomatic or highly technical combinations were excluded, following Nation's (2013) recommendation to prioritise high-frequency items.

Example selection. Concordance lines were shortened to one sentence, and lines containing unknown vocabulary or culturally inappropriate content were removed. Frankenberg-Garcia (2014) found that several carefully chosen examples support production better than a single example, so each collocation is illustrated by three to five sentences.

Classroom use

Schools in the region differ in their technical conditions, so three modes of use were envisaged.

Global Scientific Review

Full access. Learners work on the website individually or in pairs using computers or mobile phones, and additionally search SkELL directly at the Search stage.

Limited access. The teacher displays the website with a projector, and learners complete tasks on paper as a whole class or in groups.

No access. The teacher prints the tasks and concordance lines in advance. Boulton (2010) showed that such paper-based DDL materials remain effective.

In all modes, the Produce and Reflect stages are carried out in the classroom, because they require interaction and teacher feedback. The website also gives the teacher a ready source of examples for explaining errors in learners' written and spoken work.

Conclusion

"CORPUS-BASED COLLOCATIONS' LAB" was designed to bridge the gap between large research corpora and the needs of secondary school learners. Its design is based on learner fit, reduced cognitive load, guided discovery and curriculum alignment. Organising tasks according to textbook units and the six stages of the "Corpus Detective" system allows teachers to integrate corpus-based collocation work into ordinary lessons, regardless of the level of technical equipment of the school.

References

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*(3), 534–572.

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL, 21*(1), 37–54.

Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge University Press.

Davies, M. (2008–). *The Corpus of Contemporary American English (COCA)*. https://www.english-corpora.org/coca/

Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL, 26*(2), 128–146.

Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359–370). Routledge.

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal, 4*, 1–16.

Kilgarriff, A., Marcowitz, F., Smith, S., & Thomas, J. (2015). Corpora and language learning with the Sketch Engine and SKELL. *Revue française de linguistique appliquée, 20*(1), 61–80.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Global Scientific Review