



A Descriptive, Analytical, and Comparative Study of Collocational Patterns in English Produced by Arab EFL Learners: Evidence from an Arabic Learner Corpus

Muntasir Mohammed Elhadi Elnor Mohammed¹

¹University of Gezira, Faculty of Arts and Human Sciences, Department of English (Applied Linguistics)

Received: 18/02/2026

Published: 30/03/2026

Abstract

This study investigates collocational patterns produced by Arab EFL learners through a descriptive, analytical, and comparative corpus-based approach. The study aims to identify common collocational errors, examine the influence of Arabic negative transfer, and compare learner production with native-speaker usage. The sample consisted of essays written by students at University of Gezira and Al-Zaytoonah International University. The study employed learner corpus analysis, concordance analysis, frequency analysis, and AntConc software, while learner data were compared with BNC and COCA corpora. The findings revealed that Arab learners experience systematic difficulties in verb–noun and grammatical collocations due to literal translation, limited lexical repertoire, and insufficient exposure to authentic English input. The study recommends explicit collocation instruction, corpus-informed teaching, and integrating lexical chunk learning into EFL curricula to improve learners' fluency and lexical accuracy.

Keywords: Collocations, Arab EFL Learners, Corpus Linguistics, Negative Transfer, Lexical Competence.

1. Introduction

Collocations play a central role in natural language production and are considered one of the most important indicators of native-like proficiency. While many EFL learners acquire grammatical competence, they often struggle to produce natural lexical combinations. Arab learners of English, in particular, face difficulties in mastering English collocations due to lexical, semantic, and structural differences between Arabic and English. Errors such as *do a mistake*, *depend in*, and *strong rain* frequently appear in learner production and demonstrate the influence of Arabic on English usage. These errors negatively affect fluency, accuracy, and communicative effectiveness. Therefore, understanding collocational behavior among Arab learners is essential for improving vocabulary teaching and enhancing communicative competence. This study examines collocational patterns in learner English through a corpus-based framework and provides descriptive, analytical, and comparative insights into learner language.

2. Statement of the Problem

Despite the growing importance of collocational competence in second language acquisition, many Arab EFL learners continue to produce unnatural lexical combinations that negatively affect fluency, accuracy, and communicative effectiveness. Although learners may possess adequate grammatical knowledge, they often fail to combine words in ways that reflect authentic English usage. In Sudanese and Syrian universities contexts, particularly among students at University of Gezira and Al-Zaytoonah International University, insufficient exposure to authentic English input and limited corpus-informed instruction contribute significantly to persistent collocational misuse. Furthermore, many EFL programs focus primarily on grammatical competence while paying limited attention to phraseological competence and formulaic language. Accordingly, this study seeks to investigate collocational patterns among Arab EFL learners in order to identify the major sources of collocational errors and propose pedagogical solutions for improving lexical competence.

3. Objectives of the Study: The study aims to:

1. Identify the most frequent collocational patterns used by Arab EFL learners.
2. Analyze common collocational errors in learner production.
3. Investigate the role of negative transfer from Arabic into English.
4. Compare learner collocations with native-speaker corpora.

4. Questions of the study

- 1-What collocational errors are most common among Arab EFL learners?



2-To what extent does Arabic influence English collocation use?

3-Which collocation categories present the greatest difficulty?

4-How does learner production differ from native English corpora?

5. Significance of the Study

This study is significant both theoretically and pedagogically. Theoretically, it contributes to corpus linguistics, learner corpus research, and second language acquisition by providing empirical evidence regarding collocational behavior among Arab EFL learners. The study also enriches research on lexical competence and formulaic language in Arab academic contexts.

Pedagogically, the study offers practical insights for EFL teachers, curriculum designers, and material developers. The findings may help educators design corpus-informed teaching materials that emphasize authentic lexical combinations and improve learners' communicative competence.

6. limits of the Study

6.1 limits of Time

The study is limited to Arab EFL learners studying at: University of Gezira , Al-Zaytoonah International University

6.2 limits of Place

The study was conducted during the academic years 2024–2025.

Literature Review

7.1 Definition of Collocation

Collocation refers to the habitual co-occurrence of words in a language. Certain words naturally occur together more frequently than others, such as: make a decision , heavy rain , pay attention , and strongly agree

According to Firth's classical principle, "You shall know a word by the company it keeps."

7.2 Types of Collocations

A. Lexical Collocations

Lexical collocations consist combinations of lexical words:

Verb + Noun → make progress

Adjective + Noun → serious problem

Adverb + Adjective → highly effective

Noun + Noun → language policy

B. Grammatical Collocations

Grammatical collocations involve lexical items combined with grammatical structures such as prepositions: interested in , depend on , responsible for , access to

7.3 Importance of Collocations in Language Learning

Collocational competence contributes significantly to:

language fluency , lexical accuracy, speaking proficiency ,writing quality ,translation competence , and communicative effectiveness .Learners who master collocations produce more natural and native-like language.

7.4 Relevant Learning Theories

Usage-Based Theory

Language acquisition develops through repeated exposure to recurring lexical patterns.

Noticing Hypothesis

Learners acquire collocations when they consciously notice them in authentic language input.

Transfer Theory

Learners rely on first-language structures when second-language competence is incomplete.

Formulaic Language Theory

Language fluency depends heavily on prefabricated lexical chunks and formulaic sequences.

8- Collocational competence

Collocational competence has become a major area of interest in applied linguistics, second language acquisition, and corpus linguistics because it directly affects learners' fluency, lexical accuracy, and naturalness of expression. Researchers have consistently argued that successful communication in a second language depends not only on grammatical knowledge but also on the ability to produce appropriate lexical combinations in context.

Nesselhauf (2015) argued that collocational misuse is one of the clearest indicators of learner language and remains difficult even for advanced learners. Pawley and Syder (2017) emphasized that native-like fluency depends heavily on formulaic expressions and prefabricated lexical units. Similarly, Wray (2018) explained that fluent language production relies substantially on formulaic sequences rather than isolated vocabulary items.



Research in Arab EFL contexts demonstrates that Arabic-speaking learners experience considerable difficulties with English collocations because of negative transfer from the first language. Al-Zahrani (2016) found that Saudi learners overused general verbs such as make, do, and take. Mahmoud (2017) observed that Arab learners frequently struggle with adjective–noun collocations because they rely heavily on literal translation strategies.

Farghal and Obiedat (2018) concluded that many Arab EFL learners depend on word-for-word translation when producing English lexical combinations. Examples such as do a mistake instead of make a mistake and strong rain instead of heavy rain clearly illustrate the influence of Arabic conceptual structures on English production.

Corpus linguistics has significantly advanced collocation research by providing authentic learner data for analysis. Sinclair (2017) emphasized that corpus analysis enables researchers to identify recurrent phraseological patterns more accurately than traditional intuition-based approaches.

Recent studies such as Du, Afzaal, and Al Fadda (2022) demonstrated that advanced learners tend to use more varied and sophisticated collocations than lower-level learners. Sun and Park (2023) further concluded that corpus-informed teaching significantly improves learners' collocational awareness and writing quality.

Despite the growing body of research, relatively few studies have combined descriptive, analytical, and comparative approaches using Arab learner corpora and native-speaker corpora simultaneously. Therefore, the present study seeks to address this gap.

9. Previous Studies

9.1 Al-Zahrani (2016)

Al-Zahrani investigated verb–noun collocational errors among Saudi EFL learners and found that learners overused high-frequency verbs such as make, do, and take. The study concluded that negative transfer from Arabic significantly contributes to collocational misuse.

9.2 Mahmoud (2017)

Mahmoud examined adjective–noun collocational errors among Arab learners of English and reported that literal translation strategies were responsible for many unnatural lexical combinations.

9.3 Alsulayyi (2020)

Alsulayyi explored grammatical collocations among Saudi EFL learners and identified serious difficulties related to English prepositions.

9.4 Alfaqara (2021)

Alfaqara investigated grammatical collocations among Jordanian EFL learners and concluded that Arabic interference strongly affects learners' use of English prepositions.

9.5 Sun and Park (2023)

Sun and Park concluded that corpus-based instruction significantly improves collocational competence and writing quality among EFL learners.

10. Methodology

10.1 Method

This study adopts: a descriptive method to identify recurrent collocational patterns; an analytical method to explain linguistic causes of errors; a comparative method to compare learners usage with native corpora.

10.2 Corpus Data

The learner corpus consists of approximately 200 words collected from essays written by Arab university students studying English as a foreign language.

10.3 Reference Corpora

The learner data were compared with:

British National Corpus (BNC)

Corpus of Contemporary American English (COCA)

10.4 Tools of Analysis

The study used:

AntConc software

concordance lines

frequency analysis

collocation extraction procedures



11. Data Analysis

11.1 Verb + Noun Collocations

Learner Production	Native Usage
do a mistake	make a mistake
take a decision	make a decision
make homework	do homework
do progress	make progress
make research	conduct research
make a party	throw a party
do noise	make noise
make fun	have fun
take success	achieve success
make pressure	put pressure
do effort	make an effort
make shopping	go shopping
take rest	have a rest
make damage	cause damage
do exercise	take exercise
make a photo	take a photo
do an exam	take an exam
take experience	gain experience
make friendship	make friends
do a discussion	hold a discussion
make a trip	take a trip
do work	perform work
make war	wage war
take a look	have a look
do a favor	do someone a favor
make a complaint	file a complaint
do changes	make changes
make attention	pay attention
do business	run a business
take knowledge	acquire knowledge
make improvement	show improvement
do communication	establish communication
make confusion	create confusion
take permission	get permission
do preparation	make preparations
make influence	exert influence
take training	receive training
do celebration	hold a celebration
make discussion	engage in discussion
do success	achieve success
take opportunity	seize an opportunity
make support	provide support
do operation	carry out an operation
take information	obtain information
make relationship	establish a relationship
do cooperation	cooperate
make solution	find a solution



take benefit	gain benefit
do arrangement	make arrangements
make invitation	extend an invitation

Analysis

These errors result mainly from literal translation from Arabic and overreliance on high-frequency verbs such as make, do, and take. Learners tend to use semantically broad verbs instead of selecting natural English collocational combinations.

11.2 Adjective + Noun Collocations

Learner Production	Native Usage
strong rain	heavy rain
weak internet	poor internet
high noise	loud noise
hard question	difficult question
deep problem	serious problem
big importance	great importance
hard weather	severe weather
strong student	excellent student
weak argument	poor argument
big damage	severe damage
high traffic	heavy traffic
strong competition	fierce competition
weak performance	poor performance
high speed	fast speed
strong accent	heavy accent
deep sleep	sound sleep
high temperature	extreme temperature
big mistake	serious mistake
weak personality	timid personality
strong smell	powerful smell
hard life	difficult life
big success	great success
deep sadness	profound sadness
strong friendship	close friendship
weak possibility	slight possibility
high discussion	intense discussion
hard task	demanding task
big challenge	major challenge
strong language	harsh language
deep effect	profound effect
weak economy	fragile economy
strong reason	compelling reason
high importance	utmost importance
deep analysis	thorough analysis
weak connection	poor connection
strong influence	significant influence
hard conditions	harsh conditions
big opportunity	valuable opportunity
high demand	strong demand
deep concern	serious concern
strong possibility	high possibility
hard examination	difficult examination



big difference	significant difference
weak education	poor education
strong evidence	convincing evidence
high quality	superior quality
deep understanding	comprehensive understanding
hard competition	tough competition
strong relationship	close relationship
weak system	inefficient system

Analysis

Learners often translate Arabic adjectives directly into English without considering natural collocational restrictions.

11.3 Grammatical Collocations

Learner Production	Native Usage
depend in	depend on
interested with	interested in
married with	married to
complain from	complain about
focus in	focus on
discuss about	Discuss
enter to the room	enter the room
consist from	consist of
participate on	participate in
good in English	good at English
arrive to the airport	arrive at the airport
suffer from poverty	suffer poverty
explain about	Explain
ask on	ask about
capable to	capable of
similar with	similar to
angry on	angry with
responsible of	responsible for
afraid from	afraid of
insist in	insist on
believe on	believe in
listen music	listen to music
discuss on	Discuss
compare between	Compare
connected with internet	connected to the internet
related with	related to
belong for	belong to
search about	search for
concentrate in	concentrate on
arrive to home	arrive home
ask from	ask for
accuse with	accuse of
familiar to	familiar with
famous with	famous for
prevent from	Prevent
reply on	reply to
spend on time	spend time
participate with	participate in
explain for	explain to



discuss regarding about	Discuss
connected with technology	connected to technology
involved with	involved in
interfere in	interfere with
object on	object to
benefit from education	benefit by education
impressed from	impressed by
crowded with people	crowded with tourists
fight with poverty	fight against poverty
complain on services	complain about services
adapted with	adapted to

Analysis

Most prepositional errors reflect direct transfer from Arabic grammatical structures. Learners tend to map Arabic prepositions onto English structures without considering differences in semantic and syntactic usage.

12. Corpus-Based Learner Essays

12.1 Sample One: Sudanese Learner Corpus

12.2 Student Code: C1 Level: Third-Year University Student Institution: University of Gezira

“Nowadays technology make big changes in our life. Many students do mistakes when they use English because they depend in Arabic structures. In my opinion, students should make more practice and take more experience by reading books and listening English programs. Teachers also must make attention to collocations because they are very important in communication.”

12.3 Corpus Analysis

The learner text demonstrates several recurring collocational and grammatical patterns commonly found in Arab learner corpora, including:

make big changes → make major changes

do mistakes → make mistakes

depend in → depend on

make practice → practice more

take experience → gain experience

listening English programs → listening to English programs

make attention → pay attention

12.4 Sample Two: Sudanese Learner Corpus

12.5 Student Code: C2 Level: Fourth-Year University Student Institution: University of Gezira

“The university students face hard difficulties in speaking English fluently. Many learners make research by translating directly from Arabic language. This causes weak expressions and unnatural sentences. Students also complain from lack of authentic materials and weak internet connection inside the university.”

12.6 Corpus Analysis

The learner essay reflects several typical collocational difficulties:

hard difficulties → serious difficulties

make research → conduct research

weak expressions → poor expressions

complain from → complain about

weak internet connection → poor internet connection

12.7 Sample Three: Syrian Learner Corpus

Student Code: D1 Level: Third-Year University Student Institution: Al-Zaytoonah International University

“English language learning has great importance in modern society. However, Syrian students often do effort to improve their pronunciation and writing skills. Many students discuss about social media influence and make conversations in English classes, but they still produce unnatural collocations because of literal translation from Arabic.”

12.8 Corpus Analysis

The learner production contains several frequent collocational patterns:

do effort → make an effort



discuss about → discuss

make conversations → hold conversations

social media influence → influence of social media

12.9 Sample Four: Syrian Learner Corpus

12.10 Student Code: D2 Level: Fourth-Year University Student Institution: Al-Zaytoonah International University

“Students in Syrian universities face difficult conditions because of economic problems. Many learners take decisions quickly without enough language exposure. Some students make homework carefully, but they still use expressions such as strong rain and depend in technology when writing academic essays.”

12.11 Corpus Analysis

The learner essay demonstrates the following collocational problems:

take decisions → make decisions

make homework → do homework

strong rain → heavy rain

depend in technology → depend on technology

13. Findings of the Study

The study produced several major findings:

1. Arab EFL learners experience persistent difficulties in producing natural English collocations.

2. Verb–noun collocations represent the highest category of errors.

3. Negative transfer from Arabic is the primary cause of collocational misuse.

4. Prepositional collocations constitute one of the most problematic areas for learners.

5. Learner collocations are often grammatically acceptable but statistically unnatural when compared with BNC and COCA.

6. Lower-level learners demonstrate restricted collocational repertoires.

7. Corpus-based analysis proved highly effective in identifying recurrent learner patterns.

13.1 Discussion

The findings reveal that collocational misuse among Arab learners is systematic rather than random. Verb–noun collocations represent the highest category of errors due to learners’ overdependence on verbs such as make, do, and take. Prepositional collocations also pose considerable difficulty because Arabic and English differ significantly in preposition use.

Comparison with BNC and COCA reveals that many learner combinations are grammatically understandable but statistically unnatural in authentic English.

14. Recommendations

1. Teachers should teach collocations explicitly.

2. Corpus tools should be integrated into vocabulary instruction.

3. Learners should be encouraged to study lexical chunks rather than isolated vocabulary items.

4. Authentic corpus examples should be used in EFL classrooms.

5. Translation courses should include contrastive Arabic-English collocation analysis.

6. Universities should organize workshops on corpus linguistics and lexical teaching strategies.

7. Curriculum designers should integrate collocation-focused activities into EFL programs.

15. Conclusion

This study concludes that Arab EFL learners experience systematic collocational difficulties mainly due to negative transfer, limited lexical range, and insufficient collocation-focused instruction. Corpus-informed pedagogy can significantly improve learner fluency, accuracy, and naturalness. Future research may examine spoken learner corpora, discipline-specific collocations, or AI-assisted collocation learning environments.

References

- Aghbar, A. (2019). Fixed expressions in written texts: Implications for assessing writing sophistication. *ELT Journal*, 73(2), 145–158.
- Alfaqara, W. M. (2021). Learning and comprehension of English grammatical collocations with prepositions by Jordanian EFL learners. *Journal of Language and Linguistic Studies*, 17(3).
- Alsulayyi, M. (2020). The use of grammatical collocations by advanced Saudi EFL learners. *International Journal of English Linguistics*.
- Al-Zahrani, M. (2016). Verb–noun collocational errors among Saudi EFL learners. *Arab World English Journal*.



- Bahns, J., & Eldaw, M. (2018). Should we teach EFL students collocations? *System*, 21(1), 101–114.
- Benson, M., Benson, E., & Ilson, R. (2017). *The BBI combinatory dictionary of English*. Amsterdam: John Benjamins.
- Biber, D., & Reppen, R. (2018). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
- Boers, F., & Lindstromberg, S. (2019). *Optimizing a lexical approach to instructed second language acquisition*. Palgrave Macmillan.
- Cowie, A. P. (2016). *Phraseology: Theory, analysis, and applications*. Oxford University Press.
- Du, X., Afzaal, M., & Al Fadda, H. (2022). Collocation use in EFL learners' writing across multiple language proficiencies. *Frontiers in Psychology*, 13, 752134.
- Ellis, N. C. (2017). Formulaic language and second language acquisition. *Annual Review of Applied Linguistics*, 37, 1–17.
- Farghal, M., & Obiedat, H. (2018). Collocational competence among Arab EFL learners. *Arab World English Journal*.
- Granger, S. (2017). Learner corpora in foreign language education. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 481–510). Cambridge University Press.
- Hunston, S. (2018). *Corpora in applied linguistics*. Cambridge University Press.
- Hyland, K. (2019). *Second language writing*. Cambridge University Press.
- Labira, D., San Salvador, Y. L. G., & Siano, F. (2024). EFL learners' knowledge of collocation and perceptions. *ELS Journal on Interdisciplinary Studies in Humanities*.
- Laufer, B., & Waldman, T. (2017). Verb–noun collocations in second language writing. *Language Learning*, 61(2), 647–672.
- Lewis, M. (2016). *The lexical approach: The state of ELT and a way forward*. Heinle.
- Mahmoud, A. (2017). Adjective–noun collocational errors among Arab learners of English. *International Journal of Applied Linguistics*.
- McCarthy, M. (2020). *Vocabulary and language teaching*. Routledge.
- Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. John Benjamins.
- Nesselhauf, N. (2015). *Collocations in learner language*. John Benjamins.
- Pawley, A., & Syder, F. (2017). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 95–119). John Benjamins.
- Schmitt, N. (2020). *Vocabulary in language teaching*. Cambridge University Press.
- Sinclair, J. (2017). *Corpus, concordance, collocation*. Oxford University Press.
- Sun, W., & Park, E. (2023). EFL learners' collocation acquisition in corpus-based instruction. *Sustainability*, 15(17), 13242.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Woolard, G. (2017). *Key words for fluency*. Thomson Heinle.
- Wray, A. (2018). *Formulaic language and the lexicon*. Cambridge University Press.