Success Through Error: Using Error Analysis of ChatGPT Output in English as a Foreign Language Learner Writing Instruction

Arifi Waked

Prince Mohammad Bin Fahd University, Saudi Arabia, https://orcid.org/0000-0003-1222-0978

Muhammad W. Ashraf

Prince Mohammad Bin Fahd University, Saudi Arabia, https://orcid.org/0000-0002-7087-9714

Hanadi AbdelSalam

Prince Mohammad Bin Fahd University, Saudi Arabia, https://orcid.org/0000-0003-2591-4661

Khadija El Alaoui

Prince Mohammad Bin Fahd University, Saudi Arabia, https://orcid.org/0000-0002-6852-8800

Maura Pilotti

Prince Mohammad Bin Fahd University, Saudi Arabia, https://orcid.org/0000-0001-7955-680X

Abstract: Questions exist as to whether AI tools, such as ChatGPT, can aid learning. This study examined whether in-class exercises involving error detection in text generated by ChatGPT can aid students' foreign language writing. Participants were Arabic-English speakers who were classified as ranging from modest to competent English users according to IELTS standards. There were two conditions, each involving four sections of an English communication course, distributed across consecutive semesters and taught by the same instructor. In the treatment condition ($n = 101$), students performed error detection exercises using text generated by ChatGPT. In the control condition ($n = 112$), students performed the same exercises without ChatGPT. Then, they wrote a descriptive essay. It was hypothesized that if ChatGPT exercises enhanced students' confidence as writers, they would write longer essays. If ChatGPT exercises enhanced attention to detail, the exercises would increase writing quality (e.g., greater use of low-frequency and abstract words, and greater reliance on complex syntactic structures). In this study, students exposed to ChatGPT exercises used more low-frequency words and wrote more sentences but of shorter length than control condition students. The number of words did not differ, suggesting that ChatGPT exercises enhanced writers' attention to the shorter sentence length of the English language and its vast vocabulary. The results of this study suggest that ChatGPT exercises yield minor benefits for improving the writing of second-language speakers. Yet, extensive and broader exercises may yield more substantial benefits.

Keywords: Writing, English as a Foreign Language, ChatGPT, Arabic, Saudi Arabia

Introduction

In November of 2022, a controversial artificial intelligence (AI) chatbot, Chat Generative Pre-trained Transformer (ChatGPT), was released by Microsoft's OpenAL. ChatGPT uses Large Language Model (LLM) Functionality to generate written output intended to imitate human-like responses to prompts. ChatGPT has both an open access version, available online, as well as a paid version with several additional features available to subscribers. This application has gained international attention among instructors of all educational levels due to its ability to produce responses to human-initiated prompts that are accurate, informative, and lexically and semantically coherent in a wide variety of genres (Lo, 2023). As such, many educators are concerned about the possible ethical ramifications of the use of ChatGPT and other forms of LLM functionality. This is of particularly concern for instructors of writing courses, the aim of which is to improve students' compositional abilities.

When used unethically, LLM Functionality programs, such as ChatGPT, may impede student learning as well as lead to unearned, inflated grades. However, when used ethically, ChatGPT offers a variety of valuable learning opportunities to students. It can serve as a fast and reliable replacement to traditional search engines when determining answers to content-based questions. It can also be used to allow students to practice their linguistic skills by providing immediate feedback in the form of grammatical correction when their own writing is entered as a prompt (Huang et al., 2022). Indeed, ChatGPT can address a wide range of writing queries concerning syntax, semantics, language style, pragmatics, and format (Barrot, 2023). This may be of particular assistance in allowing students of writing classes the opportunity for sustained practice without fear of judgment, thereby potentially reducing writing anxiety (Beck & Levine, 2023). For educatucators teaching English-language writing to English as a foreign language (ESL) learners, it may have positive implications as a pedagogical aid within the classroom as well (AlAfnan et al., 2023). Text generated in response to prompts entered into ChatGPT can be used by writing instructors for in-class exercises such as error detection, correcting essay structure, and evaluating the appropriateness of generated content (Kohnke et al., 2023).

However, educators must also remain aware of the risk that ChatGPT may be detrimental to learning and yield serious challenges to instructors. For instance, the availability of ChatGPT may increase students' dependency on automation. This in turn may reduce their opportunities to develop, exercise, and practice critical thinking skills. More broadly, dependence on AI reduces students' development of academic agency, fostering a grade-oriented approach to learning. Students are known to display two basic and broad approaches toward their academic work (Pilotti, 2022; Pilotti et al., 2022): grade orientation (i.e., work whose goal is to receive a desired

grade) and learning orientation (i.e., work whose goal is to learn useful skills and acquire valuable knowledge). Grade orientation is recognized as counterproductive to learning and as discouraging of the adoption of advantageous study habits and ethical conduct (Daumiller & Janke, 2020; Kinne et al., 2022; Meyer et al., 2019). The danger of unwarranted dependency on automation intensifies when writing assignments do not require personalized answers that rely on critical thinking skills or do not incorporate particular lectures and textbook materials. Such writing assignments are easily completed by LLM Functionality programs, increasing the temptation of this form of plagiarism. For such assignents, it may be challenging for instructors to differentiate work produced by a student independently from work produced by AI (AlAfnan et al., 2023; Rudolph et al., 2023). As such, some academic institutions have temporarily banned students from using ChatGPT and other AI-based algorithms (Barrot, 2023). However, given the similarity of ChatGPT output to human writing, this can be difficult to enforce.

The extant literature on the possible uses of ChatGPT and other AI-powered tools is growing in two parallel directions. One developing genre of literature includes commentaries which describe the features of particular AI-powered tools. In these commentaries, reasoned arguments are often used to illustrate the potential benefits or pitfalls of each feature or set of features (e.g., Hong, 2023; Imran & Almusharraf, 2023). Empirical studies on the impact of specific uses of AI-powered tools on writing have also begun to emerge. These studies have provided evidence regarding the relative benefits, limitations, and pitfalls of these tools (e.g., Athanassopoulos et al., 2023; Baskara, 2023; Yan, 2023). For instance, Athanassopoulos et al. (2023) examined whether students' reliance on ChatGPT could improve vocabulary and grammar in German as a foreign language writing. In their study, students first were asked to write a message to a friend describing their favorite sport. They were then shown a grammatically improved version of their message with context-appropriate vocabulary as generated by ChatGPT. After reviewing this output, students were asked to write another message on a similar topic. Athanassopoulos et al. found that this exercise not only increased the length of the texts written by students (as measured by the number of words), but also the length of sentences in this writing, as is consistent with the syntactical structure of standard German writing. This exercise was also found to improve text quality as measured by the use of 'unique' words. These findings suggest that particular modes of use of ChatGPT can positively impact both the quantity and the quality of students' writing in a foreign language. Namely, notwithstanding its potentially nefarious uses, ChatGPT can aid the learning of a second language, including serving as a platform for practicing writing and offering timely suggestions for improvement.

However, when using ChatGPT as a pedagogical tool, it is important to remember that artificial intelligence is not comparable to human intelligence. LLM functionality programs retrieve information exclusively from what is available online. As such, these programs are prone to producing written output with the types of errors commonly made in human writing. For native speakers of English, learning to writing in their native language is a complex skill requiring a high working memory load along with repeated rounds of revision and error correction. However, with time and practice, native speakers of English may be able to detect the errors in ChatGPT output. This allows students to interact with the output in a way that not only improves their proofreading abilities but may also reinforce their own knowledge of the complex rules of English-language

ICRES

International Conference on
Research in Education and Science

www.icres.net | April 27-30, 2024 | Antalya, Turkiye | www.istes.org

istes
Organization

writing. As with more traditional exercises in error analysis, this may assist writers in producing future writing of their own that is free of the errors they have detected (Karr, 2022). Writing in one's second or foreign language is even more complex due to factors such as age of language acquisition, age at which one first began writing and potential transfer effects from one's native language. This may make EFL learners even more susceptible to the temptation to rely upon AI tools such as ChatGPT. However, when EFL learners rely on AI tools such as ChatGPT, they are potentially less likely to have the ability to accurately assess ChatGPT output for errors and inaccuracies than their peers for whom English is their first language.

For EFL students, one issue of concern that may impact writing in one's second language is the similarity or difference of the learner's first language. This may impact the degree to which first language interference may complicate developing English-language writing skills. Diab (1996) described the phenomenon of language interference as the detrimental role that the innate knowledge of one's native language has on the ability to learn a second or foreign language. Due to unique differences between the morphosyntactic, orthogonal, and phonological structures of English and Arabic, specific types of linguistic transfer may be more commonly found in the English-language writing of native speakers of Arabic. Modern Standard Arabic (MSA) is the high level of the diglossic linguistic systems existing within the Arabic-speaking nations of the Middle East and North Africa. Diglossia may be defined as a linguistic environment in which a language exists in two distinct forms, one viewed as the higher-level form and the other as the lower-level form. In the situation of Arabic, the lower level form is the regional dialect of colloquial Arabic which is acquired naturally and used in informal situations. The higher-level form, MSA, is standardized throughout Arabic-speaking nations and is used in formal situations. Education in Arabic-speaking nations often takes place primarily in MSA with English language instruction introduced at varying levels of education (Daqar et al., 2018).

One key difference between MSA and English is that the base form of a root word in Arabic is often a verb consisting of three consonants which can be inflected to produce nouns as well as other verbs reflecting person, tense, number, and gender (Scott and Tucker, 1974). Differences in word order and the rules of subject-verb agreement may lead to some of the common errors found in the writing of native Arabic-speaking second language learners of English. Sentences written in MSA often begin with a verb and may only state a noun or pronoun once in a paragraph. As written MSA allows subject drop, this feature often crosses over into the English-language writing of EFL learners who are native speakers of Arabic. The comma (,) also plays a notably different role in writing in MSA as compared to its role in English-language writing. This form of punctuation often functions in the capacity with which the period (.) is used in English. As such, it is not uncommon for paragraphs written in MSA to be comprised of what a native speaker of English perceives to be a single sentence with main ideas separated by commas alone. All of these grammatical features of MSA may be found in the writing of native Arabic speaking students when writing in English (Khatter, 2019).

Grammatical errors found by Diab in his 1996 study of the writing of EFL learners who were native speakers of Arabic living in Lebanon included errors in subject-verb agreement, the use of appropriate modifiers in relation to singular and plural nouns, accurate use of prepositions, and the appropriate inflection of nouns to produce

IICRES

International Conference on
Research in Education and Science

istes
Organization

www.icres.net   April 27-30, 2024   Antalya, Turkiye   www.istes.org

both plural nouns as well as their associated adjectives and adverbs. For native speakers of Arabic living in Saudi Arabia, Alkhudiry and Al-Ahad (2020) found that writing errors primarily occurred in subject-verb agreement, verb inflection, sentence and word structure, spelling irregularities, and vocabulary. Khattar (2019) also found that the majority of errors in the English-language writing in Saudi Arabian EFL students were found in punctuation. All of these areas differ greatly between MSA and English, and, as such, are prone to appear as part of the transfer effects in the English-language writing of students who have primarily been educated in MSA.

The present study examined the impact of error analysis of ChatGPT output on English-language writing for EFL native speakers of Arabic. Participants included freshman students in an English-language introductory academic writing class. Participants proficiency in written English ranged from moderate to competent. As such, this study did not focus on writing errors, but rather on the quality and quantity of various factors in participants' English-language writing. The amount written (as indexed by the number of words, number of sentences, and their length) was intended to gauge their fluency (Maisa, 2018). The quality of their writing (as indexed by the use of low-frequency words, abstract words, complex syntactic structures, and coherent content) was intended to gauge the complexity of their writing (Johnson, 2017). The in-class ChatGPT output error analysis exercise aimed to help students identify and correct errors in their own writing. These were thought to potentially include grammatical form and structure, correct selection of vocabulary to convey intended meaning, and proper essay structure and format.

This study began in the fourth week of two consecutive semesters. It began immediately prior to the assignment of a descriptive essay on a topic of the students' choice. This was the students' first assigned essay in an introductory-level academic writing course. Instruction in the first three weeks of class focused on the form and structure of the three primary sections of a standard five-paragraph essay in English (the introduction paragraph, body paragraphs, and the conclusion paragraph) as well explicit instruction on morphosyntactic rules that differ between writing in English and MSA. Prior to beginning this assignment, the study began with analysis of an instructor-generated prompt for ChatGPT to write a descriptive essay in the standard five paragraphs structure. The prompt used was "Write a five-paragraph descriptive essay on haloumi cheese. Include an introduction, three body paragraphs, and a conclusion." Haloumi cheese is a popular Middle Eastern food with which all students were familiar. Students were then provided the output in the form of a Microsoft Word document and given 15 minutes to read the output and correct errors based on their both their own knowledge and information previously learned in class. The students and instructor then engaged in a class discussion in which students shared errors they had identified in the ChatGPT-generated output. Each error identified was then discussed as a class as to how it should be corrected. The types of errors discussed primarily included grammatical and semantic errors, errors in lexical cohesion, errors in the connectivity between consecutive paragraphs, and errors in the structure and form of the essay produced. The instructor also directed students' attention to specific areas of the output produced by the program and assisted students' detection and correction of errors that students had not identified. In her instruction, the particular strengths and limitations of ChatGPT as a "writer" were underscored and students' writing competence, as indexed by their sense of agency, was emphasized. In the

control condition, students were exposed to a similar practice prior to beginning their own assigned descriptive essay, but the text examined was intentionally generated by the instructor herself with a variety of intentional errors. These including grammatical and semantic errors, errors in lexical cohesion, errors in the connectivity between consecutive paragraphs, and errors in the structure and form of the essay. As per the ChatGPT condition, the class first engaged in a discussion of their own error detection and ways in which these errors could be corrected. The instructor then assisted in error identification and correction, and in the recognition of correct forms and formats of errors that students had not identified in their independent work.

This field study was guided by two assumptions about the impact of these particular ChatGPT exercises on students' foreign-language writing, one entailing their confidence in the process of writing, and the other involving their attention to the products of their writing. Specifically, the comparison between the treatment and the control conditions gave rise to two complementary hypotheses:

H1   If ChatGPT-driven error detection exercises enhanced students' confidence as writers, students would write longer essays (i.e., the number of words and sentences used) than those of the control condition.

H2   If ChatGPT-driven error detection exercises enhanced students' attention to their writing, the exercises would enhance the quality of the essays produced by students compared to those of the control condition. Aspects of writing affected would be the use of low-frequency (rare) and more abstract words, as well as the use of complex syntactic structures and consideration of text cohesion (quality measures).

Method
Sample and Sampling

Participants included 213 freshmen enrolled in one of eight sections of a communication course devoted to EFL academic writing. Students were freshmen who were native speakers of Arabic and foreign-language learners of English. Participants ranged in age from 18 to 25 years. To enroll in the course, students had to obtain at least a 6.0 total score and a 5.5 writing score (range: 0-9) in the International English Language Testing System (IELTS). Thus, the selected participants included EFL learners who were classified as ranging from modest to competent users of the English language. The course aimed to teach freshmen how to write in academic English by offering practice leading to compositions in four genera that are clear, concise, and semantically, syntactically, and structurally accurate.

Prior to beginning this study, participants were given a three-weeks period of class focused on the form and structure of the three primary sections of a standard five-paragraph essay in English (the introduction paragraph, body paragraphs, and the conclusion paragraph) as well explicit instruction on morphosyntactic rules that differ between writing in English and MSA. In the fourth week of the semester, four sections of this course received ChatGPT error-detection exercises (treatment condition: $n = 101$), and four sections served as the control

condition (n = 112). To ensure that information about the ChatGPT error-detection exercises did not spread to the control condition, two consecutive semesters were selected. The four sections assigned to the control condition were situated in the first semester and the four sections assigned to the treatment condition were situated in the ensuing semester. The course was taught by the same instructor with 17 years of university-level teaching experience. The instructor was bilingual in English and Arabic and had been exposed to both languages since birth. Institutional records suggested that the student samples in these two semesters were demographically similar, including the number of hours in which they were enrolled [$F(1, 211) < 2.29$, ns]. The sample included only students who provided informed consent and completed all phases of the study as described below. Excluded from the sample were 29 students (11.65%) who did not complete the assignment or whose output was judged by SafeAssign as being more than 15% similar to other sources, including institutional databases, global databases, and the Internet. If content that was suspected to potentially be AI generated was found in more than 40% of students' essays, their submission would also receive a score of 0 unless the student was able to provide evidence that the work submitted was in fact their own writing. This percentage was selected due to the high instances of false identification of AI generated content among many AI detection programs (Sadasivian et al., 2023).

Materials and Procedure

The error-detection exercises used in all class sections included in this study were based on the premise that one key aspect of writing practice is helping students identify and correct errors in written communications. These exercises will in turn assist students in avoiding such errors in their ensuing writing. Such errors may include grammatical structure, meaning, and proper essay structure and form. The ChatGPT class exercises began with an instructor-generated prompt for ChatGPT version 3.0 to write a descriptive essay in the standard five-paragraph essay format. The prompt used was "Write a five-paragraph descriptive essay on haloumi cheese. Include an introduction, three body paragraphs, and a conclusion." Students were then given 15 minutes to read the output and identify errors. Following this, a class discussion ensued in which students discussed how to best correct the problems identified. The instructor then alerted students to further errors which had not yet been discussed and assisted the students in determining how they should best be corrected. Errors examined primarily included grammatical and semantic mistakes, errors in lexical cohesion, errors in the connectivity between consecutive paragraphs, and errors in the structure and form of the output produced. The particular limitations of ChatGPT as a "writer" were also highlighted. Conversely, students' agency in creating the content of their own writing was heavily emphasized. In the control condition, students examined text generated by the instructor with a variety of errors, including grammatical and semantic mistakes, errors in lexical cohesion, errors in the connectivity between consecutive paragraphs, and errors in the structure and form of the essay. As per the ChatGPT condition, through students' independent work, class discussion, and the assistance and guidance of the instructor, errors were identified, corrections were made, and attention was directed by the instructor to additional errors that had not been identified by students. Although comparisons could not be made between AI generated output and students' agency as writers in the control condition, students were praised for their abilities to identify these errors and were reminded of their agency in producing well-written compositions as EFL learners.

Beginning in the class immediately following this exercise in error analysis, after giving informed consent, participants wrote a descriptive essay on a topic of their choosing using the standard five-paragraphs essay format. They were directly informed that their work would be checked for plagiarism through SafeAssign as well as for AI generated content through an AI detection program. If their writing was found to be more than 15% similar to other sources, including institutional databases, global databases, and the Internet, it would be returned to them with a score of 0. If content that was suspected to potentially be AI generation was found in more than 40% of students' essays, their submission would also receive a score of 0 unless the student was able to provide evidence that the work submitted was in fact their own writing. This percentage was selected due to the high instances of false identification of AI generated content among many AI detection programs (Sadasivian et al., 2023).

Data Analyses

The descriptive essay written by each of the 213 students was submitted to TextEvaluator® software (Chen & Sheehan, 2015; Sheehan et al., 2014) developed by the Educational Testing Services (ETS). This software calculated a set of indices that measured variables relating to both the quantity and the quality of these written products. These measures were submitted to one-way Analysis of Variance (ANOVA). Condition served as the between-subjects factor (the control group and the group taught using ChatGPT).

Quantity measures included a count of the number of words and sentences in an essay, as well as the mean number of words per sentence. Quality measures were organized into those at the word level, and those at the sentence level. All scores ranged from 1 to 100. At the word level, quality measures included a word unfamiliarity tally, which reflected the extent to which the writer used low-frequency words, and a word concreteness tally, which reflected the writer's use of concrete words. A low score for unfamiliar words and a high score for concreteness indicated low-quality text. Word unfamiliarity was based on two corpora. One of these consisted of 400 million words developed by ETS, and the other consisted of 17 million words developed by Zeno et al. (1995). Concreteness ratings used a method developed by Coltheart (1981). These were used to compute an index of concreteness (i.e., the property of a word whose meaning can be experienced through one's senses or actions). Words that refer to concrete concepts have a behavioral advantage, which allows them to be cognitively processed more quickly and accurately than abstract concepts (Connell & Lynott, 2012).

At the sentence level, quality measures included a syntactic complexity score and a lexical cohesion score, all of which ranged from 1 to 100. The syntactic complexity score tallied a series of indices, such as the mean length of sentences, the mean number of modifiers per noun phrase, the mean number of dependent clauses per sentence, and a measure of sentence depth. The score reflected the memory load imposed by sentences with varying syntactic structures (Yngve, 1960). The lexical cohesion score indicated the extent to which a textual product provides a coherent message rather than an array of unrelated sentences. It also measured word repetition across contiguous sentences and the use of connectives. Thus, a low syntactic complexity score and a

high coherence score would indicate a text that is less complex and easy to read.

Results

Table 1 illustrates the mean of each quality and quantity measure along with the standard error of the mean. Students exposed to ChatGPT exercises wrote a greater number of sentences [$F(1, 211) = 5.26$, $MSE = 119.15$, $p = 0.023$, *partial eta^2* = 0.024] but of shorter length [$F(1, 211) = 4.63$, $MSE = 78.24$, $p = 0.033$, *partial eta^2* = 0.021] as compared to students assigned to the control condition. Students exposed to ChatGPT exercises also used more low-frequency words than students in the control condition. No other values reached significance [$Fs \leq 3.71$].

Table 1. Descriptive Statistics

Variables	Control Condition		ChatGPT Condition		Sign.
	M	*SEM*	*M*	*SEM*	
Quantity Measures					
No. of words	466.464	14.691	491.911	15.470	
No. of sentences	25.813	1.031	**29.248**	1.086	*
No. words per sentence	**20.207**	.836	17.596	.880	*
Quality Measures					
Word Unfamiliarity	54.446	1.337	**64.178**	1.408	*
Word Concreteness	48.616	1.050	47.822	1.106	
Syntactic Complexity	61.250	1.045	58.327	1.101	
Lexical Cohesion	44.964	1.023	45.950	1.078	

Note: Significant group differences ($p < .05$) are marked with an asterisk. The values in bold characters signify greater values.

Discussion

This study examined the impact of student-centered error analysis of ChatGPT output on the writing of EFL students completing their first writing assignment in an introductory-level academic writing class. This exercise led EFL students to exhibit two advantageous features in their writing. First, they segmented their writing into shorter sentences as compared to students in the control group. This led to an increased number of sentences in students' writing, and thus to text that more closely approximated English writing conventions. Second, they adopted a greater number of low-frequency words, thereby improving the complexity (quality) of their written products. These findings, which were obtained through a between-subjects design and from a homogenous sample of native speakers of Arabic, suggest that exercises using ChatGPT output can assist EFL students in

produce writing that is more closely aligned with the conventions of English (e.g., shorter sentences in English).

Our findings parallel those of Athanassopoulos et al. (2023) who reported that the feedback of ChatGPT on students' who are foreign-language learners of German led them to write longer sentences, as is consistent with German syntactical structure. This study differs from ours in the homogeneity of participants as those in their study had a variety of linguistic backgrounds. Furthermore, the impact of ChatGPT was measured through a within-subjects design (i.e., assessment before and after the use of ChatGPT-based exercises). However, the observed changes in students' writing indicated that ChatGPT-based exercises encouraged students to write sentences that more closely approximated the conventions of students' foreign language (Cooper, 1976). Similar to our findings, in which students in the testing condition were found to use an increased amount of low-frequency words, Athanassopoulos et al. also reported an increase in students' use of "unique" words following exercises utilizing ChatGPT. Although the exercises used by Athanassopoulos et al. were different from ours, both improved foreign-language learners' attention to their writing, leading to the generation of more conceptually complex words and sentences that more closely aligned with the syntactic rules of their foreign language.

Conclusion

The results of the present study are promising as they indicate that ChatGPT can potentially serve as a pedagogical tool for improving students' attention to specific aspects of their writing, thereby offering a supportive learning environment (Yan, 2023). More extensive exercises with ChatGPT output may be able to impact key writing qualities of EFL learners, such as sentence complexity and coherence. Whether exposing learners to ChatGPT output (as per our study) is a more or less effective educational strategy than asking ChatGPT to improve student-produced text (as per Athanassopoulos et al., 2023) is a matter to be investigated. The specific limitations of our study are to be considered in future research. They include a homogenous sample of foreign-language learners with a writing competency above that of novices and a first language which might present challenges that differ from those faced by native speakers of Arabic. The breadth of the exercises and their duration of use are other limitations that are to be examined as independent variables in future research.

Recommendations

Technological innovations cannot be ignored or discarded by educators. Irrespective of the opinion that an educator might have regarding ChatGPT and other forms of LLM Functionality, such AI tools will be used by students. As such, educators must work to foster an environment in which the ethical uses of AI as tools that can enrich students' learning opportunities is emphasized. Educators must demonstrate how various forms of AI can be used to help students critically analyze their academic tasks and products, including their writing. In order to discourage the use of AI tools as shortcuts for unpopular academic activities (Kim et al., 2023), educators need to emphasize the advantages of students' adoption of a learning orientation rather than a grade orientation

approach to education (Pilotti et al., 2022).

Acknowledgment

References

Abdelaal, N. (2020). *Translation between English and Arabic: A textbook for translation students and educators*. Springer Nature. https://doi.org/10.1007/978-3-030-34332-3

AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). Chatgpt as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology, 3*(2), 60-68. https://doi.org/10.37965/jait.2023.0184

Alkhudiry, R., & Al-ahad, A. (2020). Analyzing EFL discourse of Saudi EFL learners: Identifying mother tongue interference. *The Asian ESP Journal. 16*(2.1). 89-109.

Alobathani, J. O., Almalki, N. A., & Ibrahim, A. (2019). Differences between English and Arabic body idioms. *Journal of Applied Linguistics and Language Research, 6*(3), 139-143. https://www.jallr.com/index.php/JALLR/article/view/1025

Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research*, *3*(2), 818-824. https://doi.org/10.25082/AMLER.2023.02.009

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing, 57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Baskara, R. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *7*(2), 343-358. https://eric.ed.gov/?id=EJ1391490

Beck, S. W., & Levine, S. R. (2023). Backtalk: ChatGPT: A powerful technology tool for writing instruction. *Phi Delta Kappan, 105*(1), 66-67. https://doi.org/10.1177/00317217231197487

Chen, J., & Sheehan, K. M. (2015). Analyzing and comparing reading stimulus materials across the TOEFL® family of assessments. *ETS Research Report Series, 2015*(1), 1-12. https://doi.org/10.1002/ets2.12055

Cooper, T. C. (1976). Measuring written syntactic patterns of second language learners of German. *The Journal of Educational Research*, *69*(5), 176-183. https://doi.org/10.1080/00220671.1976.10884868

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A*, 497-505. https://doi.org/10.1080/14640748108400805

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition, 125*(3), 452-465.

https://doi.org/10.1016/j.cognition.2012.07.010

Daqar, R. (2018). Diglossic aphasia and the adaptation of the Bilingual Aphasia Test to Palestinian Arabic and Modern Standard Arabic. *Journal of Neurolinguistics.47.*131-144.

Daumiller, M., & Janke, S. (2020). Effects of performance goals and social norms on academic dishonesty in a test. *British Journal of Educational Psychology, 90*(2), 537-559. https://doi.org/10.1111/bjep.12310

Diab, N. (1996). The transfer of Arabic in the English writings of Lebanese students. *The ESP, Sao Paulo. 18*(1). 71-83.

Elhadary, T. (2023). Linguistic and Cultural Differences Between English and Arabic Languages and their Impact on the Translation Process. *International Journal of Language and Translation Research, 3*(2), 103-117. https://doi.org/10.22034/IJLTR.2023.175143

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, *5*(1), 37-45. https://jeti.thewsu.org/index.php/cieti/article/view/103/64

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. J*ournal of Computer Assisted Learning, 38*(1), 237-257. https://doi.org/10.1111/jcal.12610

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology, 15*(4), ep464. https://doi.org/10.30935/cedtech/13605

Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing, 37*, 13-38. https://doi.org/10.1016/j.jslw.2017.06.001

Karr, J. & White, A.(2022). Academic self-efficacy and cognitive strategy use in college students with and without depression and anxiety. *Journal of American College of Health.* DOI: https://doi.org/10.1080/07448481.2022.2076561

Khatter, S. (2019). An analysis of the most common essay writing errors among EFL Saudi female learners (Majamaah University). *Arab World English Journal. 11*(3). DOI: 10.24093/awej/vol11no3.26

Kim, S., Shim, J., & Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System, 10*(1), 79-88. https://doi.org/10.33851/JMIS.2023.10.1.79

Kinne, B. L., Goehring, M. T., & Williams, B. L. (2022). Academic entitlement and its potential educational consequences: A scoping review. *Journal of Physical Therapy Education, 36*(2), 115-121. https://doi.org/10.1097/JTE.0000000000000231

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, *54*(2), 537-550. https://doi.org/10.1177/00336882231162868

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), 410.

Maisa, S. (2018). Language Fluency and Its Evaluation: A Review of the Existing Literature. *Language in India, 18*(12), 320-329.

http://www.languageinindia.com/dec2018/sridharmaisalanguagefluencyreviewfinal.pdf

Meyer, K., Homa, N., & Marley, K. (2019). Learning and Grade Orientation in Undergraduate Students. *Midwest Journal of Undergraduate Research (MJUR)*, *10*, 15-29. https://research.monm.edu/mjur/files/2019/05/MJUR-i10-2019-2-Meyer.pdf

Pilotti, M. (2022). Adaptation to Challenging Demands: Do Individual Differences Matter? *The International Journal of Learning in Higher Education, 30*(1), 141-157. https://doi.org/10.18848/2327-7955/CGP/v30i01/141-157

Pilotti, M., Al Kuhayli, H., & Libdeh, N. A. (2022). Grade orientation, decision-making habits, and academic performance. *The International Journal of Learning in Higher Education, 29*(2), 125-140. https://doi.org/10.18848/2327-7955/CGP/v29i02/125-140

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 1-22. https://doi.org/10.37074/jalt.2023.6.1.9

Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can ai-generated text be reliably detected?. *arXiv preprint arXiv:2303.11156*.

Scott, M., & Tucker, G. (1974). Error analysis and English-language strategies of Arab students. *Language Learning. 24*(1).

Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *The Elementary School Journal, 115*(2), 184-209. https://doi.org/10.1086/678294

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*. Advanced online publication. https://doi.org/10.1007/s10639-023-11742-4

Yngve, V.H. (1960). A model and hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104, 444-466. https://dspace.mit.edu/bitstream/handle/1721.1/4453/RLE-TR-369-06411244.pdf

Zeno, S. M., Ivens, S. H., Millard, R. T., Duvvuri, R. (1995). *The educator's word frequency guide.* Touchstone Applied Science Associates.