Density and Readability of English Writing Among Arab Learners of English

Abdallah Abu Quba: King Faisal University. Foreign Language Centre. KSA.
aabuquba@kfu.edu.sa

Mohammed Nour Abu Guba: The University of Sharjah. Language Institute. United Arab
Emirates mabu-gub@sharjah.ac.ae

Ghassan Adnan Hasan: King Faisal University. ghasan@kfu.edu.sa

Correspondence author: **Abdallah Abu Qub'a** **aabuquba@kfu.edu.sa**

Abstract

Earlier research on writing problems among Arab learners of English focused on surface-level correctness. Deep-level aspects such as lexical density and text readability have been completely ignored. This study attempts to examine the readability and lexical density of English essays produced by Arab learners of English at a tertiary level. Data were collected from 100 Arab students majoring in engineering at King Faisal University. Electronic forms of the essays were analyzed and compared with essays written by English native speakers using readability and lexical density online formulas. Results show that the readability levels of Arab learners' written essays were much below the college level with shorter sentence lengths than in most published writings. Lexical density was close to the minimum score of general written texts with more glue words, more sticky sentences, more monosyllabic words, and with less variety in sentence starts. Furthermore, it was found that lexical density and readability correlated positively with perceptions of writing fluency, sophistication, and academic level. This study concludes with some pedagogical practices to increase the proficiency levels of academic writing at the college level. Further studies that compare students with different first language backgrounds at different proficiency levels and examine the extent to which grammatical errors, spelling, and punctuation errors affect text readability and lexical density at college level are highly recommended.

Keywords: academic writing; lexical density; readability; English as foreign language

1. INTRODUCTION

Academic writing is a demanding task not only to English as a foreign language (EFL) learners but also to English native speakers (e.g., Abu Qub'a, Abu Guba & Fareh, 2023). To produce a well-written academic English text, learners need to master many things including vocabulary and its usage, grammatical structure, punctuation, and spelling. However, mastering these aspects does not ensure that learners' writing is native-like. Much written work produced by non-native speakers of English with no mistakes with respect to the four aspects mentioned above still sounds foreign (Ostler, 1987). Other features such as lexical density and text readability that render a text more natural and more native-like (Halliday, 1985) are also important.

Lexical density relates to the number of lexical/content words divided by the total number of all words (Ure, 1971, cited in Nasseri & Thompson, 2021; Halliday, 1985: 64). Lexical/content words refer to nouns, verbs, adjectives, and adverbs while grammatical words (glue words) refer to pronouns, determiners, prepositions, and conjunctions. A text with a higher percentage of content words has a higher lexical density as it contains more information, which helps in clarifying the intended meaning of an academic text. Unlike spoken texts, written texts should have a lexical density at a minimum of 40%, which should be much higher for academic written texts (see Halliday, 1985). On the other hand, using a high percentage of grammatical words results in sticky sentences. Sticky sentences are sentences with a high percentage of grammatical words (which belong to the 200 most common words in English). These grammatical words are counted as empty spaces in writing and using a high proportion of them enforces readers to pass through more space to understand the actual meaning; conversely, using a low number of grammatical words in sentences makes the meaning clearer, and the reader's job easier (see https://prowritingaid.com//grammar/1008131/).

On the other hand, text readability relates to how much a text is easy to understand due to the writing style (Dubay, 2004). Readability is measured by certain formulas that count the average sentence length and the average number of syllables per word (see Methodology). Sentence length in academic writing is usually between 11 and 18 words. If the average sentence length is above 18, the writing will be felt to be too verbose and complicated; if it is below 11, then the writing is likely to be choppy (see https://prowritingaid.com//grammar/1008131/Readability).

A closely related aspect to readability is sentence variety, including sentence starts. This type of variety is essential to maintain the readers' interest, as writers manipulate variety to add emphasis, suspense, and so forth; otherwise, writing will be repetitive and dull and may confuse readers (Deveci, 2019).

Two more constructs that will be analyzed in this study are writing fluency (being readable, clear, cohesive, and easily understood) and sophistication (adding complexity and depth to the essay whilst writing concisely and clearly). In this study we will attempt to find out whether an essay's lexical density and readability will affect how English native speakers perceive the essays' fluency and sophistication besides the academic level of the writer (primary, elementary, secondary, undergraduate, or graduate).

The importance of lexical density and readability analysis is threefold: first, it gives instructors a more accurate picture of learners' lexical progress and helps teachers to reflect on their teaching and their materials. By analyzing academic essays quantitatively, teachers will have a more accurate picture of lexical progress (Webb & Paribakht, 2015) and this will help teachers modify their teaching methods and evaluate the appropriateness of their teaching materials. Second, it benefits students in developing and evaluating their own writing. Goodfellow, Lamy and Jones (2002:1) contend that lexical density "could be a first step in the development of a viable aid for learners evaluating their own writing." Finally, raising awareness of both teachers and learners of lexical density and readability will help learners use more content words and variety in their sentence structure, making their writing more enjoyable and more informative and thus facilitating the reader's job.

2. Literature Review

A number of studies dealt with lexical density and readability indices of written texts produced by learners worldwide. However, studies in the Arabic context are lacking. Regarding lexical density in written texts, Green (2012) measured the differences in cohesion and lexical density between native and non-native speakers' writings. He found that L2 learners' writing was different from native speakers' writings in that they contained more argument overlaps with more noun phrase repetition and extended pronoun references, more high-frequency content words, fewer abstract verb hyponyms with more abstract verbs than abstract nouns and less causal content.

Kondal (2015) investigated the effects of lexical density and lexical variety in language performance and proficiency of the EFL tenth grade learners' writings at a school in India. He found that learners' writing lacked rich and varied lexis, and learners tended to overuse and repeat frequent words in English especially nouns and verbs, which resulted in poor writing. Likewise, Ha (2019) conducted a study to explore how 35 undergraduate Korean students used lexical richness in written argumentative essays. He found that lexical diversity, complexity, and fluency contributed to a higher quality of written essays and students used different lexis in essays according to their linguistic competence and vocabulary knowledge.

Concerning readability levels, Gyasi (2011) analyzed the readability of 280 essays written by fresh students from four programs of study at Ghanaian universities. Using Flesch-Kincaid Reading Ease Formula, the researcher found that 55% of the essays were generally difficult to read and comprehend, which means that fresh students could produce essays at a college level. Tajbakhsh (2017) compared the readability levels in written materials by pre-intermediate English learners with and without lexico-semantic errors to identify the most frequent lexical errors and their effect on the readability level. The study found that correcting the lexico-semantic errors slightly improved the readability level of the text.

Similarly, Dakhi and Hutabarat (2018) examined the grammatical complexity and lexical density of 42 undergraduate theses in an Indonesian university using Flesch's readability analysis. They found that the average lexical density was 42.14 and the grammatical complexity was 14.54, with differences across chapters of the theses. In addition, the lexical density of the theses was acceptable while grammatical complexity was much below that written by a sixth-grade American intelligent student.

Studies investigating the role of lexical density and readability in perceptions and reactions to the quality of writing are scarce. Many previous studies concerning readability focused on reading, history, investment, or commercial science textbooks readability analysis (Apler, 2019; Gyasi & Slippe, 2019; Seyedian, Walters, & Olsavsky, 2019; Seyedian, Fitzpatrick, & Yi 2021; Hakim, Setyaningsih, & Cahyaningrum, 2021).

Turning to studies on Arab learners of English, no study, as far as the researchers know, has dealt with readability and lexical density of Arab learners' written work. Earlier research in the Arabic context focused on semantic, syntactic, punctuation and capitalization errors (e.g.,

Atashia & Al-Bahri, 2018; Hamed, 2018; Al-Ghabra, 2019; Khatter, 2019; Hussain, 2020; Musa 2021 ;). Their results pointed out that Arab learners face difficulties producing well-written texts and they attributed these difficulties mainly to L1 interference. Another more relevant study was conducted by Alhojailan (2021), who explored Saudi students' perceptions of academic writing and found that Saudi students were not aware of the characteristics of English academic writing.

This study therefore will be the first study in the Arabic context that focuses on measuring readability indices and lexical density. This paper aims to explore the characteristics of academic essays written by Arab learners at King Faisal University in terms of lexical density, and readability by comparing them with essays written by English native speakers. It is an exploratory study that will pave the way for more studies that investigate these neglected aspects and their effect on writing in English as a foreign language.

More specifically, this study tries to answer the following questions:

1. What is the percentage of lexical density used in XXX University students' essays? What is the percentage of sticky sentences? How do they compare with those produced by English native speakers?

2. How readable are the essays of XXX University students measured in terms of readability indices? What are the average sentence length and average number of syllables per word? How varied are the sentences produced by these learners? How similar are they to those produced by English native speakers?

3. Is there any correlation between lexical density and readability and perceptions of writing fluency, sophistication, and academic level?

3. Methodology

Data was collected from four sections (two male and two female) at the College of Engineering at King Faisal University during the first semester of the academic year 2019/2020. 100 Arab students were asked to write a full academic essay with about 250 words as a final exam. The essays then were analyzed in terms of their readability scores, variety, and lexical density. The topics given in the essays were general in nature. The writing was done inside the classroom using Microsoft Word under the supervision of the first researcher, and 90 minutes were allotted for the writing. The students did not know that their essays will be analyzed. Upon submitting their exam, the students were requested to sign a consent form that explained the use and storage of the final

exam paper, if they accepted their work to be used for research purposes. No student objected to that. All students had studied English for 20 hours weekly for a year before joining their college, and they did an introductory course in academic writing before participating in this study. As a result, it was expected that their essays would have high readability scores, adequate variety, and rich lexical density.

To measure the lexical density of the essays, the electronic forms of essays were analyzed using the online text analyzer software http://www.analyzemywriting.com/index.html.To measure text readability, the readability website (https://readabilityformulas.com/freetests/six-readability-formulas.php), which has seven readability formulas, was used. These formulas calculate the average sentence length, the average number of syllables per word, the average grade level, the reading age, and text difficulty (see the website for more details on the seven readability formulas). Note that one of these formulas was the widely used Flesch-Kincaid Reading Ease Score (Table 1). This score is a number between 0-100 indicating how difficult a text is to read. The higher the number is, the less difficult it is to read the text. These numbers correspond to US school grades. For example, a number between 90 and 100 corresponds to 5th grade, which means it is very easy to read and easily understood by an average 11-year-old student.

Table 1

Flesch-Kincaid Reading Ease Score

Score	School Level	Ease of readability
90-100	5th grade	Very easy to read. Easily understood by an average 11-year-old student.
80-90	6th grade	Easy to read. Conversational English for consumers.
70-80	7th grade	Fairly easy to read.
60-70	8th & 9th grade	Plain English. Easily understood by 13- to 15-year-old students.
50-60	10th to 12th grade	Fairly difficult to read.
30-50	College	Difficult to read.
10-30	College graduate	Very difficult to read. Best understood by university graduates.
0-10	Professional	Extremely difficult to read. Best understood by university graduates

To analyze sentence length and variety and sentence starts in students' written essays, the researchers used the prowritingaid website available at https://prowritingaid.com/en/Analysis/WebEditor/Go?id=1134944 and compared them with published academic writing in the same genre.

In addition, 10 band 9 IELTS essays were chosen from the IELTS blog as a reference point and were analyzed adopting the same methodology. These IELTS essays were given as model answers for band 9 in the academic writing module https://www.ielts-blog.com/.

To find out the correlation between lexical density and readability on the one hand, and perceptions of writing fluency, sophistication, and academic level of written essays on the other hand, we asked 10 English native speakers to rate 40 essays on a 10-point scale for fluency and sophistication, with 1 being the lowest and 10 the highest. For the academic level, they were asked to circle one of the five options that best fit the essay (primary, elementary, secondary, undergraduate, or graduate). The first 30 essays relate to the essays written by the study students, of which the first ten represented the highest-scoring essays on readability and lexical density, the second ten essays had medium scores, and the third ten represented the lowest-scoring essays. The last ten essays were the IELTS essays written by English native speakers.

4. Findings and Discussion

In this section, we present the characteristics of King Faisal University students' essays in terms of lexical density, readability, and sentence variety, respectively, and we show how they are different from model essays written by English native speakers. We then report on the statistical correlation between density and readability, and English native speakers' reactions with respect to fluency, sophistication, and academic level.

4.1 Lexical density analysis: what is the percentage of lexical density in King Faisal University students' essays?

Results show that there were clear differences between the essays written by students at XXX University and the 10 IELTS model essays, as shown in Table 2.

Table 2

Parts of speech analysis

Part of speech	XXX students	IELTS band 9
Nouns	31.13 %	26.32%
Adjectives	5.3 %	11.11%
Verbs	6.62 %	12.87%
Adverbs	1.32 %	5.26%
Prepositions	7.95 %	16.08%
Pronouns	11.92 %	5.26%
Auxiliary verbs	9.27 %	5.85%
Lexical density	44.37 %	55.56%

As Table 2 shows, the lexical density of King Faisal University students' writing was about 11% lower than that of the IELTS essays, with an average of 44.37 %. This percentage is close to the minimum percentage of the lexical density of written general texts. Note that lexical density of academic texts is expected to be higher than in general texts. This means that students' essays are very poor with respect to the extent of vocabulary used. The lower percentage of students' essays lexical density can be attributed to the low percentage of verbs, adjectives and adverbs and the high percentage of pronouns, and auxiliary verbs, which results in a large number of sticky sentences with a high percentage of glue words (55.73 %).

Examples of sticky sentences written by students are shown in 1. (The sentences are reproduced, with some mistakes, as they were written by the students.)

(1). King Faisal University students' sticky sentences (glue words are underlined in the sentences)

 a. The TV is very helpful for many things. (Glue words make up 62.5% of the sentence.)

 b. The TV is very important for the new babies because it is helping them to grow up and to learn some new things. (The percentage of glue words is 60.8.)

COMMON GROUND

c. There are many kids (sic) love some comedian people, and the comedian people are like a role model for them. (The percentage of glue words is 57.9.) Note here that 'like' is a preposition and therefore it is a glue word.

d. For example, there are *many* people who change their life because they want to see the bright side. (The percentage of glue words is 55.5.)

e. There are many causes to watch them and here are some of the causes. (The percentage of glue words is 71.4.5)

The low percentage of verbs among King Faisal University students could be related to Arabic language interference (L1 interference) where many sentences in Arabic are well-written without verbs. This means that students are still drawing on their first language writing mechanisms and structures, which results in poor lexical density, when transferred to English. This L1 interference is consistent with earlier findings that reported that first language interference is widespread among Arab learners of English (e.g., Abu Guba, Mashaqba, Jarbou, & Hajeid (2023), see also Abu Guba & Abu Qub'a, (2021) for Arabic rhetoric and culture interference). Another noticeable source for the overuse of glues words in King Faisal students' essays relates to the overuse of existential 'there'. It was noticed that King Faisal students tended to overuse existential 'there' most probably as a type of L1 interference where this construction is very widespread in Arabic, which means that students should be give alternative English constructions to improve their writing quality.

The percentages of glue words in students' essays seem to suggest that students' lexical resources are rather limited. Therefore, more attention should be given to the acquisition of lexical words especially verbs to enable students to acquire a richer vocabulary, which will result in higher lexical density that would improve students' writing quality (Gómez, Sotomayor, Bedwell, Domínguez, & Jéldrez, 2017; Ha,2019; Erarslan, 2021). Special attention should be given to academic word lists that are freely available on the Internet.

4.2. Readability analysis: how readable are the essays of King Faisal University students measured in terms of readability indices?

Results show that the reading level of the 100 essays according to the Flesch Reading Ease was 88.6, which is easy to read and is equal to texts written by sixth graders (11-year-old readers). This is much below the college level, which is below 50. In contrast, the readability index of Band

9 IELTS model essays was well-advanced with a score of 39.2, i.e. they are difficult to read texts at grade level 14 and a reader's age at 21-22 years (graduate college level). Table 3 gives more details.

Table 3

Readability analysis of students' essays and IELTS essays.

	King Faisal students	**IELTS Band 9**
Average sentence length	11.3	22.6
Median sentence length	11	21.6
Standard deviation of sentence length	4.7	7.2
Average number of syllables per word	1	2
Percentage of single syllables in text	72%	58%
Percentage of double syllables in text	17%	21%
Percentage of 3+syllables in text	11%	21%
Readability score: Flesch Kincaid	88.6	39.2

As can be seen from Table 3, the students' essays have noticeable issues with respect to readability. The main reasons behind that relate to sentence length and word length. Sentence length in King Faisal University students' essays was approximately 11.3 words. This corresponds to the minimum average sentence length in most published writings, which is between 11 and 18 words, and it is much lower than that of the IELTS model essays (22.6).

In addition, King Faisal University students tended to use more monosyllabic words (72%), but fewer polysyllabic ones, especially trisyllabic, quadrisyllabic, and above words. IELTS essays used more than double the percentage of trisyllabic and above words. All in all, the shorter sentence length and the shorter words used by King Faisal University students rendered the readability of their essays below the required college level. These findings show that students need to use longer sentences and more polysyllabic words to improve the readability of their writing.

Using larger numbers of words in sentences and using more complex and compound sentences would help students achieve better readability scores that would lead to a better writing quality (Ha, 2020). Students should be encouraged to write longer sentences but avoid the undesirable effects of too long sentences. In fact, sentence length is an important feature of good writing as it affects the reader's experience. Essays with many long sentences can be exhausting to

readers whereas texts with lots of short sentences can create a choppy and disjointed reading experience.

As for sentence variety, the results show that King Faisal University students did not use a variety of sentence starts. King Faisal University students used slightly less varied sentence starts than those used in published written texts. The most important characteristics of students' essays with respect to sentence variety were as follows:

- 57% of the sentences started with a subject (compared with 61% in published writing), as in 'TV now is taking a big role in our life'. (Note that the illustrative examples given here come from King Faisal University students' essays)

- 26% of the sentences started with a subordinate conjunction such as when and if (compared with 17% in published writing), as in 'when they do their hours quickly, they get a job before the other students'.

- 2% of the sentences started with an adverb such as actually and obviously (compared with 9% in published writing), as in 'consequently, the comedy shows are very useful, and helpful for everyone'.

- 5% of the sentences started with a gerund (compared with 1% in published writing), as in 'watching comedy shows can be very helpful to the students'.

- 0 % of the sentences started with coordinating and correlative conjunctions such as 'so' and 'not only but also' (compared with 9% in published writing).

Results pertaining to sentence variety show that King Faisal University students' writing tended to lack academic English native-like sentence variety. This seems to suggest that King Faisal University students are not aware of the various ways that can start English sentences. Students therefore should be guided to use more adverbs and coordinating and correlative conjunctions to start their sentences but fewer gerunds and subordinating conjunctions. Taken together, these results suggest that students have limited knowledge about using academic language in writing and their writing seems like spoken language transferred into written forms, which affects their performance and proficiency. It seems students are not aware of the peculiarities of academic English writing, which is clearly different from spoken English where informal language is tolerated, and shorter words and sentences are more common. English learners should be more aware of these characteristics to improve their writing skills.

To sum up, analysis of lexical density and readability shows that students' writing has a lower lexical density than that of English native speakers and the readability scores of their essays are considerably different from those written at college level. In the next section, we explore the effect of these aspects on English native speakers' reactions toward students' writing.

4.3 English native speakers' reactions to lexical density and readability: Is there any correlation between lexical density and readability and perceptions of writing fluency, sophistication, and academic level?

The ratings of the 10 English native speakers of the 40 essays showed positive reactions to better levels of lexical density and readability. As expected, the English native speakers' essays, which had better reliability and density scores, received more positive reactions than those written by King Faisal University students. Mean ratings of fluency, sophistication, and academic level are presented in Table 4 below.

Table 4

Mean ratings of fluency, sophistication, and academic level

	Lexical density	Readability	Fluency	Sophistication	Academic level
King Faisal University students' lowest-score essays	40	92.6	4.2	2.7	2.2
King Faisal University students' medium-score essays	45.3	88.2	4.5	3.8	2.9
King Faisal University students highest-score essays	49.8	80.6	5.4	5.5	3.4
Native speakers' essays	55.6	39.2	7.8	8.4	4.3

As can be seen from Table 4, the higher the lexical density of the essays, the more fluent, sophisticated, and the higher the academic level they were felt to be, with a small difference between lowest and medium-score essays, though. Likewise, the lower the readability score (which means more readable), the more positive the reactions were. For the academic level, the lowest-score essays were felt to belong to an elementary level, the medium-score essays to almost a

secondary level, the highest-score essays to a secondary level but the native speakers' essays were felt to belong to an undergraduate level. Clear differences were attested between the native speakers' essays and the Arab learners' essays. Although one cannot be fully certain that these reactions were driven by the essays' lexical density and readability, as other language factors do affect these reactions, the effects of density and readability cannot be ruled out. It seems that they play a role in perceptions of fluency, sophistication, and the academic level of the writer. Below, we provide statistical evidence to the extent to which these constructs are correlated.

To assess the linear relationship between readability and lexical density on the one hand, fluency, sophistication, and academic level on the other hand, a Pearson correlation coefficient was conducted. Prior to that, Cronbach's Alpha was run to check the internal consistency of the responses. Alpha stood at .86, meaning that the responses were reliable. Table 5 below presents the correlation coefficients between these constructs.

Table 5
Correlation coefficients

	Fluency	**Sophistication**	**Academic level**
Lexical density	.80**	.89**	.69**
Readability	-.90**	-.91**	-.62**
Fluency		.90**	.71**
Sophistication			.70**

**. Correlation is significant at the 0.01 level (2-tailed).

As can be seen from Table 5, very high positive correlations were attested between lexical density and fluency, sophistication, and academic level, p <.001. That points out that the higher the lexical density of the essays, the more fluent, the more sophisticated and the higher academic level they seem to be perceived to be by English native speakers. As for readability, there were very high negative (rather than positive) correlations with fluency, sophistication, and academic level, p <.001. That suggests that the higher the readability score (recall that the higher the score, the easier the writing is), the less fluent, the less sophisticated and the lower the academic level the essays were felt to be (see Appendix for details on correlation coefficients). The strong correlations between better lexical density and readability, and perceptions of fluency, sophistication, and academic level suggest that an essay's lexical density and readability could

improve the quality of academic essays. This awaits further, more controlled large-scale research in the future.

To sum up, results of the analyses of lexical density and readability besides the statistical correlations seem to suggest that the qualitative analysis of the written essays using these online applications can be a worthwhile effort. These tools could differentiate between written texts produced by non-native speakers and those written by English native speakers. Evidence for this contention comes from the results of other studies such as those of Magulod, Capulso, Tabiolo, and Ramada (2020), who reported that using such online tools could improve students writing quality and overall proficiency in academic writing. Including such tools in evaluating students' written work could be beneficial to students (Syarif & Putri, 2018), and could make their writing sound more nativelike, especially if these analyses were conducted by students themselves, particularly because these tools are easily accessible and free of charge.

5. Conclusion, Implications, and Future Research

It has been found that King Faisal University students' essays differ from IELTS model essays and published academic writing in terms of lexical density, readability, and sentence variety. Lexical density was close to the minimum score of general written texts, let alone academic ones. This was mainly caused by the low percentage of verbs, adjectives and adverbs and the high percentage of pronouns, and auxiliary verbs. The low density of students' essays could be related to L1 interference as many Arabic sentences are well-written without verbs, besides students' overuse of existential 'there' in their essays. These percentages seem to suggest that students' lexical resources are rather limited. Therefore, more attention should be given to the acquisition of lexical words, especially verbs to enable students to acquire a richer vocabulary.

Regarding text readability and sentence variety, results showed that the reading level of the 100 students' essays according to the Flesch Reading Ease was 88.6, which is easy to read and is equal to texts produced by sixth graders. This is much below the college level, which is below 50, and below the readability index of the IELTS essays, which is 39.2. The readability low score of King Faisal students' essays was attributed to the high percentage of monosyllabic words and the low percentage of polysyllabic words in addition to the small number of words per sentence (11.3 words). The number of words per sentence is quite small in comparison with that of the IELTS

model essays (22.6) and close to the minimum in published writings. Students should pay special attention to these aspects, which would help students achieve better readability scores.

The same trend applied to sentence variety. King Faisal University students' essays tended to have less sentence variety than did academic English ones. Students should be encouraged to increase their sentence variety by using more adverbs and conjunctions to start their sentences. Taken together, these results suggest that students have limited knowledge about using academic language in writing and their writing seems like spoken language transferred into written forms, which affects their academic writing negatively.

The statistically significant correlations between lexical density and readability on the one hand, and perceptions of fluency, sophistication, and academic level, on the other hand, seem to confirm that lexical density and readability could improve the quality of academic essays—a finding that needs to be corroborated by future research.

To recapitulate, the findings of this study indicate that the use of these qualitative and quantitative online applications to analyze written essays can be a fruitful endeavor. These tools were able to differentiate essays written by English native speakers from those written by Arab learners of English. Students should be encouraged to use these tools, which may increase their awareness of such aspects and ultimately render their writing more native-like. More research is needed to examine how reliable such tools could be to assess students' written work.

In light of these results, students should be made more aware of the peculiarities of academic written English (Crossley, 2020; Nasseri & Thompson, 2022). Students should be directed to use more lexical words, especially verbs to increase their texts' lexical density and consequently make them more expressive and informative. In addition, students should minimize the number of function words. This can be achieved by giving students more practice and model paragraphs to raise their awareness of the characteristics of English academic texts (Zhang, Chen, & Li, 2021, Granados, López-Jiménez, & Lorenzo, 2022).

Students should be given practical tips and strategies to increase readability levels. These may include using sentences of different lengths, combining simple sentences to make compound and complex sentences, using various ways to start their sentences, and using more academic content words. Similarly, teachers should encourage students to use more polysyllabic words which increases the readability level of their writing.

Further research that compares students with different first language backgrounds at different proficiency levels is highly recommended. This would shed light on the role of first language transfer and developmental errors. It will show whether readability and lexical density follow developmental paths, or they continue to exhibit mother tongue interference at advanced levels. Also, further research that examines the extent to which grammatical errors, spelling, and punctuation errors affect text readability and lexical density at college level is needed. A study that examines the relationship between overall marks awarded by examiners and readability and lexical density of students' texts is highly recommended.

Acknowledgments

Conflict Of Interests

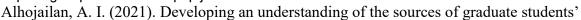The authors declare that they have no conflict of interest.

References

Abu Guba, M., & Abu Qub'a, A. (2021) Perceptions of clichés by Arab English bilinguals: Implications to academic writing. *Linguistics Journal*. 14 (2),134-154. ISSN 1718-2301

Abu Guba, M. N., Mashaqba, B., Jarbou, S., & and Hajeid. O. (2023). Production of vowel reduction among Jordanian Arab learners of English: Acoustic investigation. *Poznan Journal of Contemporary Linguistics*.

Abu Qub'a, A. Abu Guba, M. N. & Fareh, S. (2023). To what extent can Grammarly assess academic writing: The case of over-flagging. *Sage Open.*

Al-Ghabra, I.M., & Najim, A.S. (2019). Analyzing errors committed in paragraph writing by undergraduates. *Journal of Language Teaching and Research*. Vol. 10, No. 2, pp. 264-270, March 2019 DOI: http://dx.doi.org/10.17507/jltr.1002.07

Alharbi, M.A. (2019). Saudi Arabia EFL university students' voice on challenges and solution in learning academic writing. *Indonesian Journal of Applied Linguistics*. Vol. 8 No. 3. https://doi.org/10.17509/ijal.v8i3

Alhojailan, A. I. (2021). Developing an understanding of the sources of graduate students'

perceptions of academic writing. *Indonesian Journal of Applied Linguistics*, 11(2), 281-

291. https://doi.org/10.17509/ijal.v11i2.34265

Apler J. B. (2019). Readability, content, and mechanical feature analysis of selected commercial

science textbooks intended for third grade Filipino learners, *Cogent Education*, 6:1,

1706395, DOI: 10.1080/2331186X.2019.1706395

Atashian, S., & Al-Bahri, K. (2018). Towards Arab students' grammatical errors in academic

writing & their perceptions. *Arab World English Journal (AWEJ)*. DOI:
https://dx.doi.org/10.24093/awej/MEC1.10

Crossley, S.A. (2020). Linguistic features in writing quality and development: An overview. *The

Journal of Writing Research, 11*, 415-443.

Dakhi, S., & Hutabarat, H. (2018). Language effectiveness and factors influencing scientific

writing of Indonesian undergraduate thesis. English Review: *Journal of English Education*,

7(1), 61-74. doi: 10.25134/erjee.v7i1.1496.

Deveci, T. (2019). Sentence openers in academic writing: A comparison between seminar texts

and students' reflective writing papers. *Journal of Language and Linguistic Studies*,

15(1), 247-261.

Dubay, W. H. (2004). The principles of readability [Online]. Available:

http://files.eric.ed.gov/fulltext/ED490073.pdf (retrieved. Dec1 10,2020)

Erarslan, A. (2021). Correlation between metadiscourse, lexical complexity, readability and

writing performance in EFL university students' research-based essays. *Shanlax

International Journal of Education*, vol. 9, no. S1, 2021, pp. 238–254.

Granados, A., López-Jiménez, M. D., & Lorenzo, F. (2022). A longitudinal study of L2 historical writing. *Ibérica*, (43), 129–154. https://doi.org/10.17398/2340-2784.43.129

Gómez Vera, G., Sotomayor, C., Bedwell, P., Domínguez, A.M., & Jéldrez, E. (2017). Analysis of lexical quality and its relation to writing quality for 4th grade, primary school students in Chile. *Reading and Writing, 29*, 1317-1336.

Goodfellow, R., Jones, G., & Lamy, M-N (2002). Assessing learners' writing using lexical frequency. *ReCALL*, vol 14 Part 1. (133 – 145). Cambridge: Cambridge University Press.

Green, C. (2012). A computational investigation of cohesion and lexical network density in L2 writing. *English Language Teaching*, 5, 57-69.

Gyasi, W. K, (2011). An analysis of the readability of the essays of first year students of Ghanaian Universities - The Case of University of Cape Coast. *Language in India*.

Gyasi, W. K., & Slippe, D. P. (2019). Readability of English language textbooks for diploma students of the University of Cape Coast. *International Journal of Research*, 8(1), 107–115.

Ha, H. S. (2019). Lexical richness in EFL undergraduate students' academic writing. *English Teaching*, 74, 3-28.

Ha, M. J. (2020). Syntactic complexity in EFL writing: Within-genre topic and writing quality. *Computer Assisted Language Learning Electronic Journal* (CALL-EJ), 23(1), 187-205, 2022.

Hakim, A. A., Setyaningsih, E., & Cahyaningrum, D. (2021). Examining the readability level of

reading texts in English textbook for Indonesian senior high school. *Journal of English*

Language Studies, 6(1), 18. https://doi.org/10.30870/JELS.V6I1.8898

Hamed, M. (2018). Common linguistic errors among non-English major Libyan students

writing. *Arab World English Journal, 9* (3), 219-232.

Halliday, M . A . K . (1985). Spoken and written language. Geelong Vict.: Deakin University.

Hussain, M. R. A. (2020). An analysis of undergraduate Saudi EFL female students' errors in

written English essays *Arab World English Journal (AWEJ)*. Special Issue: The

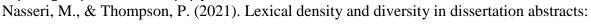Dynamics of EFL in Saudi Arabia, December 2019.

https://ssrn.com/abstract=3512533 or http://dx.doi.org/10.2139/ssrn.3512533

Khatter, S. (2019). An analysis of the most common essay writing errors among EFL Saudi

female learners (Majmaah University). *Arab World English Journal*, 10 (3) 364-381.

DOI: https://dx.doi.org/10.24093/awej/vol10no3.26

Kondal, B. (2015). 'Effects of lexical density and lexical variety in language performance and

proficiency', *International Journal of IT, Engineering and Applied Sciences Research*

(IJIEASR) 4 (10), October, pp 25–9.

Magulod, G., Capulso, L., Tabiolo, C.D., Luza, M., & Ramada, M.G. (2020). Use of technology-

based tools in ensuring quality of publishable journal articles. *International Journal of*

Learning, Teaching and Educational Research. 19. 145-162. 10.26803/ijlter.19.11.9.

Musa, F. M., (2021). The effect of grammar teaching methods on students' writing

skill. International *Journal Online of Humanities*, 7(4), 27–39.

https://doi.org/10.24113/ijohmn.v7i4.236

Nasseri, M., & Thompson, P. (2021). Lexical density and diversity in dissertation abstracts: Revisiting English L1 vs. L2 text differences. *Assessing Writing, 47*, 100511.

Ostler, S. E. (1987). English in parallels: A comparison of English and Arabic prose. In Ulla Connor & Robert B Kapalan (Eds.), Writing across languages: Analysis of L2 text. pp. 169-185. Reading, MA: Addison-Wesley.

Seyedian, M., Fitzpatrick, J., & Yi, T. D. (2021). Measuring readability levels of undergraduate investment textbooks. *Journal of Higher Education Theory and Practice*, *21*(10). https://doi.org/10.33423/jhetp.v21i10.4630

Seyedian, M., Walters, L., & Olsavsky, J. (2019). A readability analysis of undergraduate textbooks in operations management. *Journal of Business, Economics and Technology*, 22(1), 175–181.

Syarif, H & R. Putri. (2018). How lexical density reveals students' ability in writing academic text. *Lingua Didaktika*, 12(2), 86-94.

Tajbakhsh. P (2017). Lexico-semantic errors: Their effect on readability level and comprehensibility of the written texts. Volume 7, Issue 4, PP 84-88. *Journal of Research & Method in Education*.

Webb, S., & Paribakht, T.S. (2015). What is the relationship between the lexical profile of test items and performance on a standardized English proficiency test? *English for Specific Purposes*, 38, 34-43.

Zhang, H., Chen, M., & Li, X. (2021). Developmental features of lexical richness in English writings by Chinese beginner learners. *Front. Psychol.* 12:665988. doi: 10.3389/fpsyg.2021.665988

APPENDIX

Correlation coefficients

		Readability	Fluency	Lexical density	Sophistication	Academic level
Readability	Pearson Correlation	1	-.908**	-.854**	-.919**	-.621**
	Sig. (2-tailed)		<.001	<.001	<.001	<.001
	N	400	400	400	400	400
Fluency	Pearson Correlation	-.908**	1	.806**	.902**	.717**
	Sig. (2-tailed)	<.001		<.001	<.001	<.001
	N	400	400	400	400	400
Lexical density	Pearson Correlation	-.854**	.806**	1	.894**	.692**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001
	N	400	400	400	400	400
Sophistication	Pearson Correlation	-.919**	.902**	.894**	1	.704**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001
	N	400	400	400	400	400
Academic level	Pearson Correlation	-.621**	.717**	.692**	.704**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	
	N	400	400	400	400	400

**. Correlation is significant at the 0.01 level (2-tailed).