International Journal of Research Publication and Reviews

Analyzing Lexical Characteristics in IELTS Writing Task 2 Essays: A Study of Vietnamese Students from Non-English Majors

Trieu Vy Truong [a],*

[a] *Saigon International University, Ho Chi Minh City 700000, Vietnam*
E-mail address: truongtrieuvymel@siu.edu.vn

ABSTRACT:

This quantitative study examines the lexical features of IELTS Academic Writing Task 2 (IAWT2) essays written by Vietnamese non-English-major undergraduates, focusing on following writing prompts (WP): WP1 (encouraging bicycle use), WP2 (urban living and health), and WP3 (preferences for home-cooked versus fast food). A corpus of 479 essays was analyzed to explore the relationships between lexical diversity, content word types, and the use of academic and less common vocabulary. Results reveal that nouns and adjectives significantly enhance lexical diversity, while verbs and adverbs contribute minimally. Academic vocabulary coverage varied, with WP1 demonstrating the highest proportion, while WP3 relied more heavily on high-frequency vocabulary. Lexical diversity was most pronounced in WP2 and WP1, with WP3 essays showing less variation in content words. These results underscore the need for task-specific lexical training to improve students' academic writing proficiency and vocabulary use in IAWT2.

Keywords: IELTS Writing Task 2; lexical features; learner corpora; academic vocabulary; Vietnamese non-English-major students.

1. Introduction:

The IAWT2 is a key component of the IELTS exam, assessing candidates' ability to write various academic essay types, including agree/disagree, viewpoint discussion, advantage/disadvantage, and cause-solution essays (Cullen et al., 2014). These tasks require test-takers to express views on complex topics in a structured, coherent manner, mirroring academic writing demands (Nguyen, 2016). The test emphasizes flexibility and precision in vocabulary use, expecting candidates to accurately employ a wide range of lexical items with natural control and skillfully incorporate uncommon and academic vocabulary where appropriate (IELTS, 2023). However, these lexical features pose significant challenges for EFL students. Mastering a broad vocabulary for articulating complex ideas concisely is particularly challenging for lower-level learners (Ariani & Arham, 2020). The use of uncommon and academic words can also hinder the flow and sophistication of their writing (Ashraf et al., 2024), limiting non-native learners' ability to meet the high standards of the IAWT2 and potentially impacting their academic success.

Learner corpora, consisting of databases filled with texts written by learners, offer empirical evidence on the patterns of lexical features commonly adopted in their writing (Szudarski, 2018). Hence, the use of learner corpora has become an essential tool in aiding students overcome the aforementioned challenges faced by non-native and EFL students (Pérez-Paredes & Mark, 2022; Szudarski, 2018; Timmis, 2015), with the invaluable support of various corpus analysis tools of lexical features, such as Antconc (Anthony, 2024a), AntwordProfiler (Anthony, 2024b), or TAALED (Kyle et al., 2020).

The analysis of lexical features in learner corpora normally sheds light on key aspects such as lexical diversity, lexical sophistication (Vyatkina & Housen, 2020). Lexical diversity refers to the ratio between types (unique words) and tokens (total words) in a text, through the indication of Type-Token Ratio, or TTR, providing information about the range of productive vocabulary used by learners (McCarthy & Jarvis, 2007; McCarthy & Jarvis, 2010). Lexical sophistication relates to the use of advanced vocabulary that is less common and more characteristic of higher-level proficiency (Kyle & Crossley, 2014; Kyle et al., 2017). These key dimensions seem to align well with that of the requirements of lexical features IAWT2. Therefore, it is crucial to make use of this kind of learner corpora analysis in helping students tackle the challenges they face, thereby enhancing their overall language competence and performance on the test (Vyatkina & Housen, 2020).

Previous studies utilizing learner corpora to explore patterns of lexical features in student essays reveal several important findings. Ha (2019) and Maamuujav et al. (2021) both highlight the importance of lexical sophistication, noting that a higher usage of advanced vocabulary strongly correlates with better writing outcomes in academic contexts among EFL students. Conversely, Akbari (2017) finds that despite English instruction, EFL students in Ukraine struggled to diversify their vocabulary or significantly integrate academic and lower frequency words, continuing to rely heavily on high-frequency terms.

In the Vietnamese context, Nguyen et al. (2022a) and Pham & Tran (2023) report that both lexical diversity and sophistication significantly affect essay quality, with more familiar topics showing enhanced lexical traits. Meanwhile, Nguyen et al. (2022b) observed variations in lexical diversity and sophistication between two topics, though the overall band scores remained similar across these topics. These findings collectively suggest that while lexical richness and sophistication can predict higher writing scores, the influence of topics and student familiarity with content also plays a critical role.

Nevertheless, previous studies face notable limitations. Small essay sizes in learner corpora reduce statistical power, limiting generalizability and potentially skewing findings toward the traits of a few essays while overlooking broader trends (Creswell, 2018; Thompson, 2008; Biber et al., 1998). In the Vietnamese context, studies have revealed intriguing lexical patterns, but issues like small sample sizes, differences in task types, and the use of writing aids (Pham & Tran, 2023; Nguyen et al., 2022a, 2022b) raise concerns about reliability and applicability to IAWT2 essays.

Moreover, no studies have examined the relationship between the distribution of content word types (nouns, verbs, adjectives, adverbs) and lexical diversity, despite its importance for achieving higher scores in IELTS writing (Laufer & Nation, 1995; Nation, 2001; Halliday et al., 2014). Finally, while Nguyen et al. (2022a) found that Vietnamese English-major students' essays include some uncommon and academic words, it remains unclear if this applies to non-English-major students in IAWT2.

To address the identified limitations and research gaps, this quantitative study will investigate the lexical features of IAWT2 essays written by Vietnamese non-English-major students, focusing on lexical diversity, and lexical sophistication. The study draws on the corpus of 479 essays written in response to three distinct IAWT2 prompts, each representing a different task type.

The following research questions will guide this study:

1. What is the relationship between lexical diversity and the distribution of content words in the IAWT2 essays of Vietnamese non-English-major undergraduates??

2. To what extent do these students use less common and academic vocabulary in these essays?

2. Methodology:

2.1. Participants

This study used convenience sampling to recruit 180 non-English-major undergraduates from Transportation and Logistics disciplines at the university where the first author teaches. Participants, in their first or second year and classified as intermediate learners (CEFR B1-B2) based on an entrance test, underwent the Cambridge General English Proficiency Test to confirm uniform proficiency. Testing was conducted in a controlled environment, prohibiting external resources. Eleven students were excluded (eight below and three above intermediate level), leaving 169 participants (83 males, 86 females).

Participants were briefed, signed consent forms, and received 1.5 months of IAWT2 instruction in formal classes, without focus on vocabulary use. They spent about two hours weekly on writing tasks derived solely from class content. None had prior IAWT2 training, and writing prompts used in the study were withheld during instruction to minimize preparation bias and ensure spontaneous responses.

The inclusion of Vietnamese non-English-major students in this study reflects their unique challenges in mastering academic English, including syntax and vocabulary (Nguyen & Nguyen, 2022). The IAWT2 is relevant due to the IELTS exam's role in higher education and professional opportunities (Uysal, 2009). Additionally, aligning IELTS writing instruction with broader educational goals enhances their academic and occupational readiness (Trinh & Mai, 2018). This research also offers a basis for comparison with English-major students or those from different linguistic backgrounds.

2.2. IAWT2 writing prompt selection

Task selection for the IELTS IAWT2 in this study ensured relevance and fairness. A computerized system randomly chose three tasks from a database of recent IAWT2 prompts based on specific criteria. Tasks were selected to avoid prior exposure, ensuring fairness, and represented diverse task types to address various writing skills.

To validate the selections, three experienced lecturers and two writing assessment experts reviewed the prompts. They noted that some language was unclear for intermediate-level students. For example, in WP1, "Why is this the case?" was simplified to "Why?" to avoid confusion. In WP3, "traditional food" was replaced with "home-cooked food" to enhance clarity and focus.

These revisions were tested in a pilot study to confirm their effectiveness, ensuring clearer task requirements and improved student responses. This step was critical, as familiarity with topics and instructions can significantly impact written output (Nguyen et al., 2022a; Yu, 2009; Zheng, 2016)..

2.3. IAWT2 writing prompt piloting

A pilot study was conducted to validate the tasks using 19 non-English-major students at the same language level as the main study's participants. These students, not included in the main study, had received identical IAWT2 instruction. The pilot took place in the university's computer lab under real IAWT2 conditions, with each task timed at 45 minutes and no external resources allowed, completed over one-week intervals.

After each task, students completed a 6-point Likert scale survey, adapted from Malicka and Levkina (2012), on task difficulty and clarity. The survey was translated into Vietnamese and verified by experts. Additionally, five students were randomly selected for interviews to discuss the suggested task modifications.

Key observations emerged from the survey results. While participants in the pilot found task descriptions clear, , the 45-minute duration received a low mean score (M = 2.6; SD = .28) for sufficiency compared to 60-minute tasks (M = 4.1; SD = .16) with WP2 and (M = 4.2; SD = .19) with WP3. During the pilot, the 45-minute limit hindered students' ability to produce well-developed essays. Consequently, the task duration was extended to 60 minutes for the actual study, allowing participants more time to articulate their responses effectively. This change led to increased writing quality and higher token counts in WP2 and WP3 essays. Interviews revealed that four of five students found "Why?" easier to understand than "Why is this the case?" due to reduced cognitive load from abstract terms like "case" (Révész et al., 2016). Additionally, replacing "traditional food" with "home-cooked food" clarified the task by aligning better with the comparative nature of "fast food."

Table 1: Descriptive statistics of the survey

Questions	WP1*		WP2**		WP3**	
	M	SD	M	SD	M	SD
The task description of the writing prompt is easy to me	4.2	.22	4.3	.23	4.4	.31
I have enough time to do this writing prompt	2.6	.28	4.1	.16	4.2	.19
I feel relaxed to do this writing prompt	3.1	.36	4.0	.29	4.1	.30
I find the topic familiar to me	4.6	.31	4.2	.24	4.1	.25
I can easily complete this writing prompt	2.9	.13	3.9	.13	3.8	.26
** in 60 minutes						
*in 45 minutes						

Overall, task duration and simplified language significantly influenced task manageability and clarity. These adjustments ensured participants were evaluated on their analytical and writing skills rather than vocabulary interpretation (Frear & Bitchener, 2015; Johnson, 2017). Both the 60-minute duration and revised tasks were adopted for the actual study.

2.4. Data collection

Data collection followed the pilot study procedures, with each writing prompt completed within 60 minutes in university computer labs over one-week intervals, without external resources. Of 507 essays produced, 479 were included in the final corpus. Essays under 100 words (11 from WP1, 10 from WP2, and 10 from WP3) were excluded, as such brevity limits reliable corpus analysis (Anthony, 2022). The data underwent thorough cleaning, including spelling corrections to ensure accuracy and usability, as most corpus tools cannot automatically address such errors. While learner errors were corrected, their analysis was beyond this study's scope. Finally, all essays were encoded in UTF-8 format to ensure compatibility with corpus analysis tools.

Table 2: Total tokens in each writing prompt and the corpus

WP1 (N=159)	42442
WP2 (N= 160)	38179
WP3 (N= 160)	41305
Total	121926 (tokens)

Although smaller than large-scale corpora like COCA and BNC (with hundred millions of tokens), this corpus offers distinct advantages. It preserves the authenticity of naturally occurring text from IAWT2 essays, reflecting genuine student writing and capturing context-specific nuances (McEnery & Brookes, 2022). The corpus is highly representative of the target population, enabling meaningful insights into the linguistic abilities of these test-takers. Its manageable size allows for in-depth contextual analysis, closely linking findings to the essays' production context (McEnery & Brookes, 2022). Additionally, it addresses sample size limitations highlighted in prior studies (Akbari, 2017; Ha, 2019; Maamuujav et al., 2021; Nguyen et al., 2022a;

Nguyen et al., 2022b; Pham & Tran, 2023), ensuring more reliable and generalizable results. Thus, the corpus effectively meets the study's research objectives.

2.5. Data analysis

This study combines advanced software tools and custom methods for corpus analysis (Anthony, 2022). TAALED 1.4.1 (Kyle et al., 2020) was used to measure lexical diversity (TTR and TTR_CW) and content words (CW). AntWordProfiler (Anthony, 2024) assessed vocabulary coverage, focusing on academic and less common words using the New Academic Word List (NAWL) (Browne et al., 2013) and the COCA/BNC word family lists (Nation, 2004). These tools provided insights into the prevalence of high-frequency, academic, and less common vocabulary, with NAWL offering broader coverage of academic language than Coxhead's (2000) AWL. While overlaps exist between NAWL and COCA/BNC lists, only about 33% of NAWL tokens are among the 1–3k COCA/BNC words, minimizing impact on findings (Akbari, 2017; Nation, 2004).

For aspects not covered by these tools, custom Python scripts on Google Colab used nltk POS-tagging to calculate ratios of nouns, verbs, adjectives, and adverbs to total tokens in each essay. This approach was validated with a Computational Linguistics expert and tested on pilot essays, confirming reliability. By integrating advanced software and custom solutions, the study addressed previous limitations, achieving a comprehensive linguistic analysis.

Table 3: Test of normality of all variables

Note: TTR=traditional Type-Token Ratio; TTR_CW=tradional Type-Token Ratio for content words

		Kolmogorov-Smirnova		
		Statistic	df	Sig.
WP1	TTR	.045	159	.200*
	TTR_CW	.071	159	.047
	noun	.099	159	.001
	verb	.143	159	.000
	adj	.099	159	.001
	adv	.160	159	.000
WP2	TTR	.047	160	.200*
	TTR_CW	.052	160	.200*
	noun	.137	160	.000
	verb	.171	160	.000
	adj	.107	160	.000
	adv	.177	160	.000
WP3	TTR	.059	160	.200*
	TTR_CW	.043	160	.200*
	noun	.098	160	.001
	verb	.155	160	.000
	adj	.131	160	.000
	adv	.147	160	.000
*. This is a lower bound of the true significance.				
a. Lilliefors Significance Correction				

The analysis for the first question followed a similar approach, focusing on lexical diversity. We examined the relationship between content word types and two measures of diversity: traditional Type-Token Ratio (TTR) and TTR for content words (TTR_CW). For the second question, descriptive statistics were used to analyze token counts and the percentage of less common and academic vocabulary within the corpus.

3. Findings:

3.1. The relationship between lexical diversity and content words in each writing prompt

Table 4: Descriptive statistics of TTR and TTR content words in each writing prompt

Writing prompt	TTR		TTR Content words	
	M	SD	M	SD
WP1	.56	.05	.70	.06
WP2	.57	.05	.70	.06
WP3	.53	.05	.63	.07

For TTR, WP1 and WP2 demonstrated similar levels of lexical diversity (M = 0.56, SD = 0.05 and M = 0.57, SD = 0.05, respectively), while WP3 exhibited the lowest lexical diversity (M = 0.53, SD = 0.05). The TTR for content words was identical for WP1 and WP2 (M = 0.70, SD = 0.06), whereas WP3 showed reduced diversity (M = 0.63, SD = 0.07). These findings reveal that WP3 elicited less diverse lexical choices compared to WP1 and WP2, particularly for content words.

Table 5: ANOVA analysis between TTR and content word types in WP1

		Sum of Squares	df	Mean Square	F	Sig.
noun	Between Groups	.151	148	.001	50.959	.000
	Within Groups	.000	10	.000		
	Total	.151	158			
verb	Between Groups	.072	148	.000	3.747	.013
	Within Groups	.001	10	.000		
	Total	.073	158			
adj	Between Groups	.104	148	.001	3.307	.020
	Within Groups	.002	10	.000		
	Total	.107	158			
adv	Between Groups	.047	148	.000	16.010	.000
	Within Groups	.000	10	.000		
	Total	.048	158			

Table 6: ANOVA analysis between TTR content words and content word types in WP1

		Sum of Squares	df	Mean Square	F	Sig.
noun	Between Groups	.143	143	.001	1.875	.082
	Within Groups	.008	15	.001		
	Total	.151	158			
verb	Between Groups	.071	143	.000	3.103	.008
	Within Groups	.002	15	.000		
	Total	.073	158			
adj	Between Groups	.105	143	.001	5.392	.000
	Within Groups	.002	15	.000		
	Total	.107	158			

adv	Between Groups	.044	143	.000	1.321	.277
	Within Groups	.004	15	.000		
	Total	.048	158			

The ANOVA analyses in Table 5 and Table 6 reveal notable differences between TTR and TTR content words in their associations with content word types in WP1. While TTR shows strong variability across all word types, particularly nouns ($F_{(148, 10)} = 50.96$, $p < .001$) and adverbs ($F_{(148, 10)} = 16.01$, $p < .001$), TTR content words emphasize variability in adjectives ($F_{(143, 15)} = 5.39$, $p < .001$) and verbs ($F_{(143, 15)} = 3.10$, $p = .008$). Overall, TTR captures broader lexical diversity, whereas TTR content words focus on more specific lexical patterns, especially adjectives and verbs in WP1.

Table 7: Correlation between TTR and content word types in WP1

		TTR	noun	verb	adj	adv
TTR	Pearson Correlation	1	.147	.026	.329**	.056
	Sig. (2-tailed)		.064	.744	.000	.485
	N	159	159	159	159	159
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						

Table 8: Correlation between TTR for content words and content word types in WP1

		TTR_CW	noun	verb	adj	adv
TTR_cw	Pearson Correlation	1	-.016	-.132	.296**	.034
	Sig. (2-tailed)		.841	.096	.000	.669
	N	159	159	159	159	159
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						

The correlation analyses between TTR and content word types (nouns, verbs, adjectives, and adverbs) show a moderate positive correlation was observed between TTR and adjectives, $r = .329$, $p < .001$, and between TTR_CW and adjectives, $r = .296$, $p = .001$, indicating that adjectives significantly contribute to lexical diversity. Conversely, the correlations between TTR and nouns ($r = .147$, $p = .064$) and verbs ($r = .026$, $p = .744$) were weak and not statistically significant. Similarly, the correlation between TTR_CW and verbs ($r = -0.132$, $p = .096$) and adverbs ($r = -0.070$, $p = .380$) were also weak and non-significant. These results suggest that adjectives are key drivers of lexical diversity, while verbs and adverbs show little influence in WP1, especially when focusing on content words.

Table 9: ANOVA analysis between TTR and content word types in WP2

		Sum of Squares	df	Mean Square	F	Sig.
noun	Between Groups	.128	142	.001	61.267	.000
	Within Groups	.000	17	.000		
	Total	.128	159			
verb	Between Groups	.047	142	.000	4.501	.000
	Within Groups	.001	17	.000		
	Total	.048	159			
adj	Between Groups	.067	142	.000	39.878	.000
	Within Groups	.000	17	.000		
	Total	.067	159			
adv	Between Groups	.028	142	.000	1.992	.051

Within Groups	.002	17	.000	
Total	.030	159		

Table 10: ANOVA analysis between TTR for content words and content word types in WP2

		Sum of Squares	df	Mean Square	F	Sig.
noun	Between Groups	.120	137	.001	2.429	.009
	Within Groups	.008	22	.000		
	Total	.128	159			
verb	Between Groups	.047	137	.000	4.948	.000
	Within Groups	.002	22	.000		
	Total	.048	159			
adj	Between Groups	.063	137	.000	2.980	.002
	Within Groups	.003	22	.000		
	Total	.067	159			
adv	Between Groups	.029	137	.000	3.691	.000
	Within Groups	.001	22	.000		
	Total	.030	159			

In Table 9, nouns show the largest contribution to TTR (F = 61.27, p < .001), followed by adjectives (F = 39.88, p < .001) and verbs (F = 4.50, p = .000), with adverbs showing marginal significance (F = 1.99, p = .051) in WP2. In contrast, Table 10 highlights verbs (F = 4.95, p < .001) and adverbs (F = 3.69, p < .001) as stronger contributors to TTR content words, with nouns (F = 2.43, p = .009) and adjectives (F = 2.98, p = .002) playing relatively smaller roles in WP2. Therefore, nouns dominate overall lexical diversity in TTR, while verbs and adverbs gain prominence in the context of content words.

Table 11: Correlation between TTR and content word types in WP2

		TTR	noun	verb	adj	adv
TTR	Pearson Correlation	1	.243**	-.136	.118	-.158*
	Sig. (2-tailed)		.002	.087	.137	.046
	N	160	160	160	160	160
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						

Table 12: Correlation between TTR for content words and content word types in WP2

		TTR_CW	noun	verb	adj	adv
TTR_cw	Pearson Correlation	1	-.003	-.047	-.087	-.030
	Sig. (2-tailed)		.973	.551	.271	.707
	N	160	160	160	160	160
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						

For TTR in WP2, nouns showed a moderate positive correlation (r = .243, p = .002), while adjectives (r = .118, p = .137) and adverbs (r = -.158, p = .048) had weak relationships. Verbs showed a weak negative correlation (r = -.138, p = .087) with TTR. In contrast, TTR_CW analysis revealed a strong negative correlation with verbs (r = -.583, p < .001) and a moderate positive correlation with adjectives (r = .311, p < .001). Nouns (r = -.003, p = .973)

and adverbs (r = -.030, p = .669) showed negligible relationships with TTR_CW. These findings suggest that TTR emphasizes nouns, while TTR_CW highlights the roles of adjectives and verbs in lexical diversity for WP2.

Table 13: ANOVA analysis between TTR and content word types in WP3

		Sum of Squares	df	Mean Square	F	Sig.
noun	Between Groups	.141	144	.001	2.984	.009
	Within Groups	.005	15	.000		
	Total	.146	159			
verb	Between Groups	.064	144	.000	2.355	.030
	Within Groups	.003	15	.000		
	Total	.067	159			
adj	Between Groups	.064	144	.000	1.053	.488
	Within Groups	.006	15	.000		
	Total	.070	159			
adv	Between Groups	.037	144	.000	2.120	.049
	Within Groups	.002	15	.000		
	Total	.039	159			

Table 14: ANOVA analysis between TTR for content words and content word types in WP3

		Sum of Squares	df	Mean Square	F	Sig.
noun	Between Groups	.131	141	.001	1.100	.431
	Within Groups	.015	18	.001		
	Total	.146	159			
verb	Between Groups	.061	141	.000	1.235	.314
	Within Groups	.006	18	.000		
	Total	.067	159			
adj	Between Groups	.067	141	.000	2.701	.009
	Within Groups	.003	18	.000		
	Total	.070	159			
adv	Between Groups	.035	141	.000	1.170	.367
	Within Groups	.004	18	.000		
	Total	.039	159			

In Table 13, nouns (F = 2.98, p = .009) and verbs (F = 2.36, p = .030) showed significant contributions to TTR, while adjectives (F = 1.05, p = .488) and adverbs (F = 2.12, p = .049) had weaker impacts. In contrast, Table 14 revealed that adjectives were significant contributors to TTR_CW (F = 2.70, p = .009), whereas nouns (F = 1.10, p = .431), verbs (F = 1.23, p = .314), and adverbs (F = 1.17, p = .367) were not. As a result, nouns and verbs drive overall lexical diversity (TTR), while adjectives are more influential in content-word-specific analyses (TTR_CW) in WP3.

Table 15: Correlation between TTR and content word types in WP3

		TTR	noun	verb	adj	adv
TTR	Pearson Correlation	1	.243**	-.136	.118	-.158*
	Sig. (2-tailed)		.002	.087	.137	.046

	N	160	160	160	160	160
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						

Table 16: Correlation between TTR for content words and content word types in WP3

		TTR_CW	noun	verb	adj	adv
TTR_cw	Pearson Correlation	1	.145	-.121	.011	-.257**
	Sig. (2-tailed)		.068	.128	.886	.001
	N	160	160	160	160	160
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						

In Table 15, nouns showed a moderate positive correlation with TTR (r = .243, p = .002), while adverbs demonstrated a weak negative correlation (r = -.158, p = .046). Verbs (r = -.136, p = .087) and adjectives (r = .118, p = .137) had weak and non-significant relationships with TTR. In contrast, Table 16 revealed that adverbs were negatively correlated with TTR_CW (r = -.257, p = .001), while nouns (r = .145, p = .068) and verbs (r = -.121, p = .128) showed weak and non-significant relationships. Adjectives demonstrated no significant correlation with TTR_CW (r = .001, p = .886). Again, nouns drive overall lexical diversity in WP3, while adverbs negatively impact content-word-specific diversity in TTR_CW.

3.2. The extent Vietnamese non-English-major students use less common and academic vocabulary in their essays

Table 17: The coverage of NAWL and the COCA/BNC word family list

Word list	WP1 (42442 tokens)		WP2 (38179 tokens)		WP3 (41305 tokens)	
	WP1_token_coverage	WP1_%cover	WP2_token_coverage	WP2_%cover	WP3_token_coverage	WP3_%cover
NAWL	19066	44,92%	13158	34,46%	8370	20,24%
1k	27663	65,18%	27720	72,61%	33395	80,76%
2k	2748	6,47%	3615	9,47%	2767	6,69%
3k	1414	3,33%	0	0,00%	1237	2,99%
4k	1608	3,79%	414	1,08%	763	1,85%
5k	1698	4,00%	314	0,82%	202	0,49%
6k	103	0,24%	305	0,80%	195	0,47%
7k	51	0,12%	86	0,23%	50	0,12%
8k	29	0,07%	44	0,12%	32	0,08%
9k	22	0,05%	55	0,14%	64	0,15%
10k	16	0,04%	10	0,03%	9	0,02%
11k	15	0,04%	11	0,03%	11	0,03%
12k	5	0,01%	12	0,03%	4	0,01%
13k	0	0,00%	13	0,03%	4	0,01%

14k	1	0,00%	14	0,04%	2	0,00%
15k	0	0,00%	15	0,04%	0	0,00%
16k	1	0,00%	16	0,04%	4	0,01%
17k	0	0,00%	17	0,04%	1	0,00%
18k	1	0,00%	0	0,00%	0	0,00%
19k	0	0,00%	3	0,01%	3	0,01%
20k	0	0,00%	0	0,00%	0	0,00%
21k	0	0,00%	0	0,00%	0	0,00%
22k	1	0,00%	1	0,00%	0	0,00%
23k	0	0,00%	2	0,01%	0	0,00%
24k	0	0,00%	1	0,00%	0	0,00%
25k	2	0,00%	0	0,00%	0	0,00%
Out of list	249	0,59%	310	0,81%	200	0,48%

What stands out from the table is that WP1 demonstrated the highest NAWL coverage (44.82%), while WP3 showed the lowest (20.24%). Across all prompts, the 1k word band from the COCA/BNC list dominated, particularly in WP3 (80.76%). The 2k band contributed most to WP2 (9.47%), while higher bands (beyond 2k) and out-of-list tokens collectively accounted for less than 1% in all prompts. These findings suggest that WP1 relies more heavily on academic vocabulary, WP2 demonstrates moderate coverage, and WP3 uses simpler, high-frequency words. The reduced NAWL coverage in WP3 reflects a simpler lexical style with less focus on academic terms.Tokens in higher word bands (beyond 2k) contribute minimally across all writing prompts, highlighting a general reliance on more common vocabulary, not less common one. Last but not least, WP2 features the most out-of-list tokens (0.81%), suggesting greater lexical variability or creativity.

Discussion:

In terms of lexical diversity, adjectives emerged as key drivers (e.g., r = .329 with TTR in WP1), while nouns and adverbs played a smaller role. The findings align with research emphasizing varied vocabulary as a hallmark of writing quality (McCarthy & Jarvis, 2007; Afshar et al., 2017). Compared to English-major students, who often demonstrate broader lexical diversity due to higher exposure and proficiency (Nguyen et al., 2022a), non-English-major students in this study relied heavily on adjectives to expand vocabulary. This highlights a gap that targeted pedagogical interventions can address. Additionally, the focus on content-word-specific TTR adds to previous studies (Crossley & McNamara, 2010; Kyle et al., 2017), providing detailed insights into how different word types influence diversity specifically in IAWT2 essays.

The findings also reveal varied use of academic vocabulary, with WP1 showing the highest NAWL coverage (44.92%) and WP3 relying more on high-frequency words (80.76% from the 1k band). This aligns with Laufer & Nation (1995) and Nguyen et al. (2022a), who observed the challenges EFL learners face in incorporating advanced vocabulary. Compared to English-major students, who are more accustomed to using academic words, non-English-majors in this study relied on simpler vocabulary, especially for less familiar topics. These findings underscore the importance of tailoring prompts to students' knowledge levels to facilitate sophisticated lexical use (Nguyen et al., 2022b; Yu, 2009). The observation that WP1 had the highest NAWL coverage supports previous research on topic familiarity influencing lexical sophistication and provides empirical evidence specific to IAWT2 contexts.

The findings of this study have addressed the measurements important implications for both IAWT2 instruction and vocabulary teaching. For IAWT2, the results highlight the need for targeted training that addresses the lexical demands of specific task types, ensuring that students are equipped to use appropriate nouns and adjectives to enhance lexical density and diversity (Kopple, 1994; Laufer & Nation, 1995). Vocabulary instruction should focus on expanding students' range of academic and less common words, particularly those relevant to argumentative and discursive writing (Coxhead, 2000; Nguyen et al., 2022a). This can be achieved through explicit teaching of high-frequency academic vocabulary, strategies for incorporating descriptive language, and exercises that promote the effective use of advanced and varied vocabulary (Kyle & Crossley, 2015). Additionally, integrating task-specific practice, where students engage with prompts that align with IAWT2 requirements, can help bridge the gap between vocabulary knowledge and application, ultimately improving their overall writing performance (Nguyen & Nguyen, 2022).

This study has some limitations. First, the one-week interval between writing tasks may not have allowed participants sufficient time to develop their lexical skills. Future research could adopt a longitudinal learner corpus approach to track vocabulary changes over extended periods (Szudarski, 2018;

Timmis, 2015). Second, this study relied solely on quantitative analysis, which, while insightful, may overlook some nuanced aspects of lexical features. Future research should incorporate qualitative methods, such as analyzing the collocation patterns in students essay to see how content words types interact with each other.

Conclusion:

This study provides valuable insights into the lexical features of IAWT2 essays written by Vietnamese non-English-major undergraduates. Analyzing a corpus of 479 essays across three distinct prompts, the research highlights significant relationships between lexical diversity and content word types. Nouns and adjectives emerged as pivotal in enhancing lexical diversity, while verbs and adverbs contributed less prominently. The variation in academic word usage across the prompts underscores the influence of topic familiarity on lexical sophistication. WP1, focusing on encouraging bicycle use, displayed the highest academic vocabulary coverage, likely due to greater familiarity with the topic. In contrast, WP3, discussing food preferences, relied heavily on high-frequency vocabulary, indicating challenges in using advanced lexical resources for less familiar topics.

These findings emphasize the critical role of task-specific vocabulary training in preparing students for IAWT2. Targeted interventions should focus on enhancing students' ability to use academic and less common vocabulary effectively, particularly in argumentative and discursive tasks. Additionally, the study underscores the importance of providing prompts that balance linguistic challenges with student familiarity to elicit richer lexical outputs. Future research could extend these findings through longitudinal studies and incorporate qualitative analyses of collocation patterns to deepen understanding of lexical development.

References

Afshar, H. S., Movassagh, H., & Arbabi, H. R. (2017). The interrelationship among critical thinking, writing an argumentative essay in an L2 and their subskills. *Language Learning Journal*, *45*(4), 419–433. https://doi.org/10.1080/09571736.2017.1320420

Anthony, L. (2022). What can corpus software do? In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 103–125). https://doi.org/10.4324/9780367076399-9

Anthony, L. (2024a). *AntConc* (Version 4.3.1) [Software]. Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software/AntConc

Anthony, L. (2024b). *AntWordProfiler* (Version 2.2.1) [Software]. Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software/AntConc

Ariani, N., & Arham, M. (2020). An analysis of EFL test takers' problems in IELTS writing task. *Tamaddun*, *19*(2), 132–140. https://doi.org/10.33096/tamaddun.v19i2.81

Ashraf, A., Ashraf, S., & Lodhi, M. A. (2024). Identification of the challenges faced by students in IELTS writing test and designing writing practice plan. *Jahan-e-Tahqeeq*, *7*(3). https://doi.org/10.61866/jt.v7i3.1785

Baker, M., & Croft, W. (2016). Lexical categories: legacy, lacuna, and opportunity for functionalists and formalists. *Annual Review of Linguistics*, *3*(1), 179–197. https://doi.org/10.1146/annurev-linguistics-011516-034134

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Browne, C., Culligan, B., & Phillips, J. (2013). *A New Academic Word List.* http://www.newacademicwordlist.org/

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238. https://doi.org/10.2307/3587951

Creswell, J. W., & Guetterman, T. C. (2018). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson.

Crossley, S. A., & McNamara, D. S. (2010). Predicting second language writing proficiency: the roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, *35*(2), 115–135. https://doi.org/10.1111/j.1467-9817.2010.01449.x

Cullen, P., French, A., & Jakeman, V. (2014). *The official Cambridge guide to IELTS: Student's book*. Cambridge University Press.

Frear, M. W., & Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing*, *30*, 45–57. https://doi.org/10.1016/j.jslw.2015.08.009

Guo, L., Crossley, S. A., & McNamara, D. S. (2013). Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing*, *18*(3), 218–238. https://doi.org/10.1016/j.asw.2013.05.002

Ha, H. S. (2019). Lexical richness in EFL undergraduate students' academic writing. *English Teaching*, *74*(3), 3–28. https://doi.org/10.15858/engtea.74.3.201909.3

Halliday, M., Matthiessen, C. M., Halliday, M., & Matthiessen, C. (2014). An introduction to functional grammar. In *Routledge eBooks*. https://doi.org/10.4324/9780203783771

IELTS. (2023). *IELTS band descriptors*. https://ielts.org/take-a-test/your-results/ielts-scoring-in-detail

Jhangiani, R., Chiang, I. A., Cuttler, C., & Leighton, D. C. (2015). *Research methods in psychology* (2nd ed.). BC Campus.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Lund Working Papers in Linguistics*, *53*, 61–79.

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

Kopple, W. J. V. (1994). Some characteristics and functions of grammatical subjects in scientific discourse. *Written Communication*, *11*(4), 534–564. https://doi.org/10.1177/0741088394011004004

Kyle, K., & Crossley, S. A. (2014). Automatically assessing lexical sophistication: indices, tools, findings, and application. *TESOL Quarterly*, *49*(4), 757–786. https://doi.org/10.1002/tesq.194

Kyle, K., Crossley, S. A., & Jarvis, S. (2020). Assessing the validity of lexical diversity indices using direct judgements. *Language Assessment Quarterly*, *18*(2), 154–170. https://doi.org/10.1080/15434303.2020.1844205

Kyle, K., Crossley, S., & Berger, C. (2017). The tool for the automatic analysis of lexical sophistication (TAALES): Version 2.0. *Behavior Research Methods*, *50*(3), 1030–1046. https://doi.org/10.3758/s13428-017-0924-4

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*(3), 307–322. https://doi.org/10.1093/applin/16.3.307

Maamuujav, U., Olson, C. B., & Chung, H. (2021). Syntactic and lexical features of adolescent L2 students' academic writing. *Journal of Second Language Writing*, *53*, 100822. https://doi.org/10.1016/j.jslw.2021.100822

McCarthy, P. M., & Jarvis, S. (2007). vocd: A theoretical and empirical evaluation. *Language Testing*, *24*(4), 459–488. https://doi.org/10.1177/0265532207080767

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, *42*(2), 381–392. https://doi.org/10.3758/brm.42.2.381

McEnery, T., & Brookes, G. (2022). Building a written corpus: what are the basics? In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 35–47). https://doi.org/10.4324/9780367076399-4

Moore, T., & Morton, J. (2004). Dimensions of difference: a comparison of university writing and IELTS writing. *Journal of English for Academic Purposes*, *4*(1), 43–66. https://doi.org/10.1016/j.jeap.2004.02.001

Nation, P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language: Selection, Acquisition, and Testing* (pp. 3–13). https://doi.org/10.1075/lllt.10.03nat

Nguyen, B. T. T., Nguyen, T. T. B., & Phan, Q. N. (2022). Vietnamese IELTS learners' perceptions of IELTS writing task 2 and their performance-based lexical resource. In *IELTS Research Reports Online Series, No. 5/22*. British Council, Cambridge Assessment English and IDP: IELTS Australia. https://www.ielts.org/teaching-and-research/research-reports

Nguyen, H. N., & Nguyen, D. K. (2022). Vietnamese Learners' performance in the IELTS Writing Task 2. *International Journal of TESOL & Education*, *2*(1), 170–189. https://doi.org/10.54855/ijte.222111

Nguyen, T. B. T., Newton, J., & Tran, Q. P. N. (2022). The effect of topic on EFL writing by Vietnamese tertiary students: Insights from combining a lexical richness analysis with student self-reports. *Language Related Research*, *13*(5), 511–540. https://doi.org/10.52547/lrr.13.5.19

Nguyen, T. M. H. (2016). A review of IELTS writing test. *The University of Danang - Journal of Science and Technology*, *10*(107), 16–18. https://jst-ud.vn/jst-ud/article/view/1794

Panahi, R., & Mohammaditabar, M. (2015). The strengths and weaknesses of Iranian IELTS candidates in academic writing task 2. *Theory and Practice in Language Studies*, *5*(5), 957–967. https://doi.org/10.17507/tpls.0505.10

Pérez-Paredes, P., & Mark, G. (2022). What can corpora tell us about language learning? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 313–327). https://doi.org/10.4324/9780367076399-22

Pham, T. T. H., & Tran, H. P. (2023). Lexical features and the quality of Vietnamese EFL students' online English writings. *Journal of University Teaching and Learning Practice*, *20*(2). https://doi.org/10.53761/1.20.02.06

Révész, A., Kourtali, N., & Mazgutova, D. (2016). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*(1), 208–241. https://doi.org/10.1111/lang.12205

Szudarski, P. (2018). *Corpus linguistics for vocabulary: A guide for research*. Routledge.

Thompson, B. (2008). *Foundations of Behavioral Statistics: An Insight-Based Approach*. Guilford Press.

Timmis, I. (2015). *Corpus linguistics for ELT: Research and practice*. Routledge.

Trinh, T. T. H., & Mai, T. L. (2018). Current challenges in the teaching of tertiary English in Vietnam. In J. Albright (Ed.), *English tertiary rducation in Vietnam* (pp. 40–53). Routledge. https://doi.org/10.4324/9781315212098-4

Uysal, H. H. (2009). A critical review of the IELTS writing test. *ELT Journal*, *64*(3), 314–320. https://doi.org/10.1093/elt/ccp026

Vidakovic, I., & Barker, F. (2010). Use of words and multi-words unites in Skills for Life in writing examination. In *IELTS Research Report 41*. IELTS Australian/British Council.

Vyatkina, N., & Housen, A. (2020). Complexity. In N. Tracy-Ventura & M. Paquot (Eds.), *The Routledge handbook of second language acquisition and corpora*. Routledge.

West, M. (1953). *A General Service List of English words: With semantic frequencies and a supplementary word-list for the writing of popular science and technology*. Longman.

Woods, K., Hashimoto, B., & Brown, E. K. (2022). A multi-measure approach for lexical diversity in writing assessments: Considerations in measurement and timing. *Assessing Writing*, *55*, 100688. https://doi.org/10.1016/j.asw.2022.100688

Yu, G. (2009). Lexical diversity in writing and speaking task performances. *Applied Linguistics*, *31*(2), 236–259. https://doi.org/10.1093/applin/amp024

Zheng, Y. (2016). The complex, dynamic development of L2 lexical use: A longitudinal study on Chinese learners of English. *System*, *56*, 40–53. https://doi.org/10.1016/j.system.2015.11.007