HERMES

Journal of Language and Communication in Business

No. 65 – 2025

Published by the School of Communication and Culture, Aarhus University

ISSN: 1903-1785

HERMES
Journal of Language and Communication in Business

Aliasghar Kargar

Lexical Diversity and Sophistication in Professional Architectural Discourse: A Computational Comparison of Native and Non-native English Writing

Abstract

The current study aimed to explore the lexical differences between texts authored by native (NES) and non-native English-speaking (NNES) professional architects submitted to ArchDaily, the world's most visited architecture website. The study focused on lexical diversity and sophistication indices, which proved to be theoretically and operationally pertinent to L1 and L2 writing discriminations. The corpus of the study comprised randomly selected texts in the category of residential architecture as the most common instances of architecture projects written by Iranian and British architects. As stated on the website, the texts are authored and revised by the architects themselves and are strongly advised to undergo thorough review and verification for accuracy and quality. The data underwent analysis using Coh-Metrix Core Desktop Beta (2023) package, with the results subsequently input into SPSS for further analysis. Preliminary analysis revealed statistical variances between most of the diversity and sophistication indices with supplementary analysis indicating that lexical indices such as word frequency, familiarity, hypernymy, and diversity significantly contributed to discerning between NES and NNES authors' compositions. The findings indicated that although English for General Academic Purposes (EGAP) expectations and the criteria for successful academic writing remain consistent, English for Specific Academic Purposes (ESAP) and English for Specific Purposes (ESP) stylistic conventions are adaptable and evolve according to the needs of the target audience. As the non-expert readership continues to grow, academic writing in architecture tends to employ simpler, less sophisticated vocabulary to enhance accessibility and comprehension. The results highlight the significance of adhering to the specialized community of practice stylistic conventions, which carries important pedagogical implications for instructional approaches in ESP and academic writing programs.

Keywords

academic writing, Coh-Metrix, EGAP, ESAP, ESP, lexical diversity, lexical sophistication

1. Introduction

Lexical choices are believed to be a valid index of both writing quality and the writer's language proficiency (Doró & Pietilä, 2015), and when it comes to the communication practices of professionals in ESP settings, inappropriate or uncommon vocabulary items may lead to misunderstanding or even being overlooked by the community of practice. Previous studies with native English-speaking (NES) and non-native English-speaking (NNES) writers in various academic contexts confirm the existence of such differences (Silva, 1993; Pípalová, 2015), though the need for further research is still felt (Crossley & McNamara, 2009).

The issue can be more effectively illustrated by considering the contextual dimensions of English for Academic Purposes (EAP) and distinguishing its subfields: English for General Academic Purposes (EGAP) and English for Specific Academic Purposes (ESAP) (Carkin, 2005). According to Carkin, NNES authors typically learn to write within either the EGAP or ESAP contexts. EGAP encompasses fundamental academic language components and skills applicable across disciplines, whereas ESAP focuses on authentic disciplinary texts and discipline- and genre-specific language features. From a lexical perspective, since ESAP involves scientific discourse identification, under-

* *Aliasghar Kargar*
Faculty of Archeology and Museum
Department of English Teaching
Shiraz University of Arts
a_kargar@shirazartu.ac.ir

standing rhetorical organization, genre analysis, and attending to the linguistic properties unique to academic disciplines, it is assumed that NNES writers benefit more from ESAP than EGAP in acquiring what Perez-Llantada (2012) terms "standardized scientific English".

While most previous research has focused on evaluating the lexical properties of academic essays (Kyle, Crossley, & Jarvis, 2021) and publications such as scientific articles, theses, and dissertations (Lei & Yang, 2020), comparatively little attention has been paid to examining lexical variations among NES and NNES authors in professional contexts. The present study, therefore, aims to address this research gap by investigating whether there are notable lexical differences between texts authored by NES and NNES professionals published on ArchDaily, the world's most visited architecture website. To address this aim, the study adopts a quantitative, corpus-based design, computes standard indices of lexical diversity and sophistication (via Coh-Metrix) on ArchDaily project descriptions by NES and NNES architects and complements the computational analysis with a brief qualitative illustration.

This focus was motivated by the authors' previous observations of such variations in Architecture ESP classes. The NES architects were selected from UK-based firms with a top-tier reputation in the global architecture community and were renowned for their innovative designs and commitment to sustainability. Evidence from their professional practice suggests that these professionals have developed ESP knowledge, which Hyland (2016, p. 61) defines as "an ability to integrate subject matter knowledge with unfamiliar, situationally appropriate linguistic and rhetorical conventions." Conversely, NNES architects were recruited from Iran-based firms whose publicly available portfolios (at the time of sampling) indicated a predominantly domestic project orientation. Most held degrees from national universities; sectoral studies report that EGAP/ESAP provision in Iran is often limited or uneven (Kaivanpanah, Alavi, Bruce, & Hejazi, 2021). While individual overseas training or collaborations cannot be ruled out, the firms' public profiles during the sampling window suggested a mainly national orientation of practice.

Accordingly, the study aimed to identify the probable lexical deviations of NNES from the NES professionals' "register of academic writing…comprising a sub-set of lexico-grammatical features and rhetorical conventions" (Hyland, 2016, p.61). Specifically, the study investigates the following research questions:

1. Are the NES and NNES professional architects' descriptive texts different in terms of lexical indices?
2. What measures of lexical analysis have the potential to discriminate between texts written by NES and NNES professional architects?

Following this introduction the paper presents the theoretical background in Section 2, the literature review in Section 3, and the method in Section 4. Section 5 reports the results, including descriptive statistics, t-tests, logistic regression and a brief qualitative illustration. Section 6 discusses the findings and Section 7 offers conclusions, limitations and pedagogical implications.

2. Theoretical Background

Learning words of a language seems an endless task for many language learners, as even at advanced levels, they may experience some lexical gaps. In an attempt to give a practical definition of lexical knowledge, Read (2000) used the umbrella term of lexical richness, referring to the statistical measures reflecting effective vocabulary use and deeper levels of lexical knowledge in speech and writing. These measures, as outlined by the author, include lexical variation, lexical sophistication, lexical density, and the number of errors. Although the classification is still used in a number of studies, attempts have been made to eliminate shortcomings in calculations and promote the appli-

cations. However, among the measures, lexical diversity and sophistication are more frequently acknowledged as two major sub-constructs of lexical knowledge in language acquisition (Crossley & Kim, 2022). This is why these two measures were chosen as the valid indicators of NES and NNES written texts in this study.

In what follows, we treat lexical diversity and lexical sophistication as the two core sub-constructs of lexical knowledge that guide our measures and hypotheses. In particular, we use lexical diversity in the contemporary, distributional sense, a construct believed to reflect the writers' mastery of a language, as more proficient writers are assumed to use a wider variety of lexical items (Kyle, Crossley, & Jarvis, 2021). Similarly, lexical sophistication is adopted here in a dual-trait sense, in writing understood as both breadth and depth of vocabulary knowledge (Garner, Crossley, & Kyle, 2020). According to the authors, while depth refers to "how well a word is known" in terms of its number of senses (polysemy) and its specificity (hypernymy), breadth refers to the number of unique words in a text, regardless of how often they appear. Lexical diversity, on the other hand, is a more nuanced measure that considers not just the number of different words, but also the distribution of word usage within the text. More proficient writing is characterized by both depth and breadth of vocabulary items, including diversity of words and specificity of words appearing in fewer contexts (Kyle & Crossley, 2015). As a strong predictor of writing quality and language proficiency, lexical sophistication also refers to the index of text evaluation, which measures diversity, concreteness, familiarity, and specificity of lexical items (Crossley & Kyle, 2018). It is believed that more proficient writers have wider lexical knowledge and use more sophisticated words (less frequent and familiar) instead of more frequent and general words to show subtleties in meaning and express their ideas more concisely and clearly. Crossley (2020) lists some features for sophisticated lexical items as more academic, more specific, more difficult to process and recognize, longer, less concrete, less imageable, less familiar, less frequent, and less diverse based on the context, and concludes that these are the words produced by more proficient writers.

Thanks to the computational development and psychological theories, there have been plenty of indices and lexicon repertoire available for analyzing the lexical properties of a text. One of the most valid and widely used computational tools for analyzing various text lexical features is Coh-Metrix (Dowell, Graesser, & Cai, 2016). McNamara, Graesser, McCarthy, and Cai (2014) elaborate on the essential databases and computational tools incorporated in Coh-Metrix as follows. The first item, as mentioned by the authors, is the MRC Psycholinguistics Database, offering 26 psychological properties for 150,000 words. The properties include age of acquisition, degree of familiarity of a word to an adult, degree of concreteness or abstractness, imageability or the ease of creating a mental image for a word, and ratings of meaningfulness. The second database available for Coh-Metrix is CELEX Word Frequency Database. In this database, word frequency is defined as "the relative frequency of the words in public documents per million words" (p. 42). Also, word frequency is an important feature, as there is a direct relationship between word frequency, word length, and text difficulty. The next database with different linguistic and psychological features for more than 170,000 words is the WordNet database. The words in the database are arranged to show their place in their lexical network, or how they are semantically related; for example, synonymous, polysemous, and hypernymy relations are identified. Polysemy refers "to the number of senses associated with a word' and is used as "an index of text ambiguity" (McNamara et al., 2014, p. 75), whereas hypernymy refers to "the place of each word in a hierarchical scale of specificity where higher values reflect more specific words" (p. 76).

3. Literature Review

In EGAP contexts many studies have examined NES/NNES writing differences with Coh-Metrix as the main tool. A representative example is Crossley and McNamara (2009), whose corpora comprised 195 English argumentative essays by native Spanish speakers and 208 by NES undergra-

duates at Mississippi State University. The selected indices targeted lexical networks and cohesion (lexical coreferentiality, semantic coreferentiality via latent semantic analysis, word frequency and MRC-based lexical information, hypernymy and polysemy, spatiality, and causality). ANOVA analysis identified ten variables with the largest effects between groups. Results showed that NNES writers used fewer abstract and context-independent words and exhibited less developed lexical networks/cohesion, whereas NES writers displayed wider networks with greater access.

Still within EGAP, lexical sophistication and cohesion indices align closely with human ratings and proficiency levels. Previous research indicates several features of NNES texts, including fewer nominalizations, lexical control, diversity and sophistication (Jung, Crossley, & McNamara, 2019). Jung et al. (2019) conducted a study to determine which features of text structure, cohesion, lexical sophistication and syntactic complexity, as measured by Coh-Metrix, distinguish levels of NNES writing proficiency on the Michigan English Language Assessment Battery (MELAB). The corpus comprised 750 compositions (200–300 words) randomly sampled from MELAB tests in 2013, responding to 42 prompts across 21 forms; expert raters were trained on topic development, text length, appropriateness of lexical choices, morphological and syntactic control, cohesion and accuracy of spelling and punctuation (p. 40). Indices such as frequency, concreteness and familiarity (i.e., less familiar and more abstract content words) were highly predictive of writing scores, and the findings indicated that breadth and appropriateness of lexical choices characterized higher-level writing. A number of cross-sectional Coh-Metrix studies show that more frequent words are learned earlier; longitudinal work, however, reports that while frequency (at type and token level) is a function of time on task (Crossley, Skalicky, Kyle, & Monteiro, 2019), progress co-occurs with increased use of more academic, infrequent and, in some contexts, more concrete words (Crossley & Kim, 2022).

EGAP evidence also supports the validity of lexical-diversity metrics, with abundance emerging as the clearest signal and length-robust measures mitigating TTR's sensitivity to text length. In a recent review, Kyle et al. (2021) evaluated diversity indices and dimensions—abundance (number of types), variety (relative proportion of unique words), and volume (number of tokens)—using a corpus of 615 argumentative essays by NES and NNES writers on two prompts. Lexical diversity, defined as "the variety of word use that can be found in a person's speech or writing" (p. 160), was assessed by two trained raters on a 1–10 scale, and all indices were computed with the open-source Tool of Automatic Analysis of Lexical Diversity (TAALED). Abundance (number of types) correlated most strongly with human judgements; a higher number of types may signal the introduction of new words through idea generation, paraphrasing and summarizing, which is indicative of higher proficiency. Because diversity scores are length-dependent, proficient writers may receive lower scores in longer texts; accordingly, Covington and McFall's (2010) MATTR and McCarthy and Jarvis's (2010) MTLD were introduced as length-independent alternatives. Taken together across EGAP studies, cohesion/lexical-network and lexical-sophistication indices consistently track human ratings and proficiency, and diversity metrics are validated with abundance carrying the strongest signal; longitudinally, writers shift from more frequent to less frequent, more academic items.

By contrast, in ESAP/ESP the primary driver shifts from nativeness to genre enculturation. Apparent L1 advantages reflect enculturation rather than nativeness per se (Lei & Yang, 2020), and lexical behavior is discipline-/genre-sensitive (Anderson, 2014). Lei and Yang (2020) compared lexical-diversity indices across Chinese PhD candidates' research-article manuscripts, native final-year undergraduates, native master's students, and native expert authors, finding that the native experts outperformed the other groups. Because that expert cohort was entirely L1, the advantage is best read as nativeness coupled with long-term genre enculturation—i.e., an expertise/enculturation effect rather than nativeness per se. Similarly, for lexical sophistication in ESAP—and drawing on the same evidence base (Lei & Yang, 2020)—studies likewise point to genre expertise/enculturation rather than nativeness per se as the main discriminator. Overall, these studies indicate that the criterion of "nativeness as the representation of writing standards" requires revision (McKinley & Rose, 2018),

although the finding that native experts outperformed non-native experts reflects the conflation of nativeness with long-term enculturation. This ultimately suggests that, where enculturation is comparable, nativeness is unlikely to be decisive.

In addition to the body of research on EGAP contexts, multiple studies examining ESAP contexts have yielded comparable results. For example, Azadnia (2022) analyzed the statistical differences between NNESs' and NESs' Coh-Metrix indices of lexical density, diversity, and sophistication in TEFL dissertations, indicating that the NES group used higher proportions of unique, high-frequency, familiar, concrete, imageable, meaningful, and multi-sense ambiguous words. In another study investigating word length, lexical frequency, and lexical sophistication of ESP texts, Nie (2024) compared the Medical English Corpus (MEC), Marine Engineering English Corpus (MEE), and the Corpus of English for Science and Technology (JDEST). The results indicated that lexical indices are genre- and discipline-specific (e.g., higher sophistication in medical; greater diversity in science/technology).

Taken together, while the literature is replete with EGAP studies, research on ESAP lexical properties remains comparatively limited; to the best of the authors' knowledge, no studies have specifically examined the lexical characteristics of architecture-oriented texts aimed at a global mixed audience of experts and non-experts. This gap motivates the present analysis.

4. Method

This study was primarily quantitative and descriptive. It used the Coh-metrix Core Desktop Beta (2023) package indices to describe and analyze lexical differences in NS and NNS architects' writings for ArchDaily. However, a brief qualitative analysis of two NES and NNES sample texts was conducted to complement and clarify the more complicated computational findings.

4.1 Corpus

The corpus of the study comprised the descriptive texts written and submitted by the world-leading architects to ArchDaily.com, the largest architecture website with more than 13.6 million monthly visits and offices in five countries (ArchDaily, 2024). The texts, authored and revised by the architects themselves, are strongly advised to undergo thorough review to verify their accuracy and quality. The genre is characterized by a combination of informative and descriptive elements, often aimed at showcasing architectural projects or discussing industry trends. These texts typically take the form of professional project descriptions and commentary texts aimed at both industry insiders and a broader audience interested in architecture. Recognizing this genre is essential for understanding the linguistic features examined, as it influences lexical choices and rhetorical strategies specific to professional architectural discourse in an online format. As the study aimed to compare the texts written by NES and NNES professional architects, a total of 219 texts, 114 for the NES and 105 for NNES writers, were randomly selected from the texts under the category of residential architecture, the commonest instances of architecture projects, and other categories, including higher-stakes investment and construction projects were not examined in this study. The website has made it feasible to select the country/region of the architecture projects, so the native English-speaking (NES) architects were chosen from leading British firms recognized for their top-tier reputation within the global architecture community. The size of the sample texts seems to satisfy McNamara et al.'s criterion (2014) that researchers need between 20-30 thematically related, representative, and balanced text instances for each variable. After the random selection, as the Coh-metrix analysis requires clean data (McNamara et al., 2014), the texts were reexamined and edited for possible misspelling and other oddities like pictures, illustrations, odd letters, and mathematical numbers. All texts were in .txt format files and had a code to show their category.

4.2 Measures and Tools

Coh-Metrix has been established as a robust computational tool for analyzing various lexical features of texts, such as word frequency, concreteness, polysemy, word meaningfulness, hypernymy, word age-of-acquisition scores, word imageability, and word familiarity measures (Crossley & McNamara, 2009, p.122). Previous research has affirmed the tool's reliability and validity across multiple studies (e.g., Crossley & McNamara, 2009; Dowell, Graesser, & Cai, 2016; McNamara, Louwerse, McCarthy, & Graesser, 2010). For the current study, we utilized the most recent version of Coh-Metrix, the Coh-Metrix Core Desktop Beta (2023) package, to analyze the data.

4.3 Design and Statistical Analysis

In the initial phase of utilizing the Coh-Metrix tool, the selection of variables should be guided by both theoretical frameworks and practical considerations. McNamara et al. (2014) warn that using too many or inappropriate variables can threaten validity, while too few can yield inconclusive results. The common remedy is the Train and Test approach by putting about 300 items into a training set (two-thirds) and a test set (one-third) and retaining only those meeting predefined criteria for the test set. However, this study did not use Train and Test for two reasons: (1) the sample size fell short of the 300-item threshold, and (2) strong theoretical justifications can render this approach unnecessary (McNamara et al., 2014, p.168). Instead, we focused on indices that are both theoretically and operationally relevant for distinguishing L1 vs. L2 writing, lexical diversity and lexical sophistication (Crossley & Kim, 2022). The table below (Table 1) presents a compilation of these indices, along with their corresponding codes and descriptions in Coh-Metrix (McNamara et al., 2014, pp. 247-251):

	Label in Coh-Metrix	Description
	Lexical Diversity Indices	
1	LDTTRc	Lexical diversity, type-token ratio, content word lemmas
2	LDTTRa	Lexical diversity, type-token ratio, all words
3	LDMTLDa	Lexical diversity, MTLD, all words
4	LDVOCDa	Lexical diversity, VOC, all words
	Lexical Sophistication Indices	
1	WRDFRQc	CELEX word frequency for content words, mean
2	WRDFRQa	CELEX log frequency for all words, mean
3	WRDFRQmc	CELEX log minimum frequency for content words, mean
4	WRDAOAc	Age of acquisition for content words, mean
5	WRDFAMc	Familiarity for content words, mean
6	WRDCNCc	Concreteness for content words, mean
7	WRDIMGc	Imageability for content words, mean
8	WRDMEAc	Meaningfulness, Colorado norms, content words, mean
9	WRDPOLc	Polysemy for content words, mean
10	WRDHYPn	Hypernymy for nouns, mean
11	WRDHYPv	Hypernymy for verbs, mean
12	WRDHYPnv	Hypernymy for nouns and verbs, mean

Table 1. Coh-Metrix Indices of the Study

The sample texts were then fed into the Coh-metrix Core Desktop Beta (2023) package to find out which measures of lexical analysis had the potential to discriminate between texts written by NES and NNES architects, and the following results were obtained.

5. Results

5.1 Descriptive Statistics and T-Test Results

As commonly employed techniques in Coh-Metrix analysis (McNamara, et al., 2014, p.188), preliminary results included the groups' descriptive statistics, independent-samples t-test and eta squared:

	Label in Coh-Metrix	Group	N	Mean	SD	t	df	sig.	eta squared
Lexical Diversity Indices									
1	LDTTRc	Native	114	.290219	.0540559	.541	176.4	.589	
		Non-native	105	.286851	.0352114				.001
2	LDTTRa	Native	114	.512133	.0731014	-4.687	217	.000	
		Non-native	105	.553728	.0578766				.092
3	LDMTLDa	Native	114	66.583886	20.5342695	-8.969	217	.000	
		Non-native	105	92.819245	22.5816808				.271
4	LDVOCDa	Native	114	63.789914	19.2305540	-7.448	217	.000	
		Non-native	105	81.349693	15.5911407				.203
Lexical Sophistication Indices									
1	WRDFRQc	Native	114	1.976714	.1035251	6.376	217	.000	
		Non-native	105	1.892930	.0908802				.157
2	WRDFRQa	Native	114	2.973324	.1073744	9.487	199.3	.000	
		Non-native	105	2.847886	.0860891				.293
3	WRDFRQmc	Native	114	.866867	.3845943	4.471	216.8	.000	
		Non-native	105	.621605	.4271611				.084
4	WRDAOAc	Native	114	350.680534	22.5307805	1.687	217	.093	
		Non-native	105	345.600649	22.0026250				.012
5	WRDFAMc	Native	114	561.926073	8.0565316	4.140	217	.000	
		Non-native	105	557.604439	7.3920306				.073
6	WRDCNCc	Native	114	412.074009	21.0512849	-10.04	217	.000	
		Non-native	105	439.410860	19.2102795				.317
7	WRDIMGc	Native	114	435.039972	17.7604608	-9.553	217	.000	
		Non-native	105	457.529061	17.0692560				.296
8	WRDMEAc	Native	114	437.454780	12.9924420	-4.634	217	.000	
		Non-native	105	444.854016	10.5923305				.092
9	WRDPOLc	Native	114	3.852943	.3911455	-1.426	186.9	.156	
		Non-native	105	3.918956	.2798693				.009
10	WRDHYPn	Native	114	6.601400	.3299301	3.475	217	.001	
		Non-native	105	6.453956	.2979809				.053
11	WRDHYPv	Native	114	1.550448	.1846179	-4.481	217	.000	
		Non-native	105	1.659553	.1756256				.084
12	WRDHYPnv	Native	114	2.098086	.1703083	-4.143	217	.000	
		Non-native	105	2.192456	.1666539				.073

Table 2. Results of Descriptive Statistics, T-Test and Eta Squared of the Data

5.1.1 Lexical Diversity Indices

The first part of the results indicated that among the lexical diversity indices, the mean difference between the two groups' LDTTRc (type-token ratio of content word lemmas) was not statistically significant (NS: M=.290219, SD=.0540559; NNS: M=.286851, SD=.0352114; t (176.4) =.541, p=.589). For other indices, however, the mean differences were statistically significant (LDTTRa: t (217) = -4.687, p<.0005; LDMTLDa: t (217) = -8.969, p<.0005; LDVOCDa: t (217) = -7.448, p<.0005) and eta-squared results showed that the magnitudes of the differences were very high (LDTTRa: eta=.092; LDMTLDa: eta=.271; LDVOCDa: eta=. 203). In other words, the data suggest

that, based on the definition of the diversity indices in (McNamara et al., 2014, p.67), NNES writers (with higher means) introduced "more unique words (tokens) to the texts in relation to the total number of words" (LDTTRa, LDMTLDa, and LDVOCDa), but not when only content words were counted (LDTTRc).

5.1.2 Lexical Sophistication Indices

The second part of the table presents statistics related to lexical sophistication, including CELEX-based word frequency indices such as age of acquisition, familiarity, concreteness, imageability, meaningfulness, polysemy, and hypernymy. Among these metrics, statistically significant mean differences between NES and NNES writers were observed for all measures except for the age of acquisition (WRDAOAc: NS: M = 350.680534, SD = 22.5307805; NNS: M = 345.600649, SD = 22.0026250; t(217) = 1.687, p = 0.093) and polysemy (WRDPOLc: NS: M = 3.852943, SD = 0.3911455; NNS: M = 3.918956, SD = 0.2798693; t(186.9) = -1.426, p = 0.156). Large effect sizes (eta) were found for 'word frequency measures' (0.92, 0.271, 0.203, respectively), 'concreteness' (0.317), 'imageability' (0.296), and 'meaningfulness' (0.92). In contrast, moderate effect sizes were observed for 'hypernymy for verbs' (0.084), 'nouns and verbs' (0.073), and 'familiarity' (0.073). 'Hypernymy for nouns' (0.053), however, showed a small effect size.

McNamara et al. (2014, p. 74) note that age of acquisition reflects when words enter children's vocabulary. For example, "cortex," "dogma," and "matrix" (AOA ≈ 700) are older than "milk," "smile," and "pony" (AOA ≈ 202). The lack of a significant difference in AOA means-of-content words between NES and NNES writers suggests similar MRC timeline norms. In line with these findings, the lack of a significant difference in polysemy index means indicates that both groups obtained similar WordNet polysemy values for the content words in their texts.

Moderate eta squared for the familiarity index suggests that NES writers used words with slightly higher MRC familiarity ratings, indicating the use of more familiar words. Additionally, low eta squared for WRDHYPn suggests that NNES writers used less specific nouns. In contrast, NNES writers had higher means and a high magnitude of difference between the means of WRDHYPv and WRDHYPnv, indicating the use of more specific verbs by NNES writers.

All measures of word frequency obtained very high magnitude of difference, suggesting that NES writers used more frequent words. According to McNamara, et al. (2014, p.73), "words that occur with a higher frequency are more familiar to the reader and are processed more quickly". On the contrary, the significant difference of the WRDCNCs means and high magnitude of the difference suggest that NNES writers used more concrete words. For imageability and meaningfulness measures, NNES writers obtained higher means, and the magnitude of the differences was very high. A higher Imageability score means that NNES writers used more words whose mental images were easier to construct. NNESs' higher mean of WRDMEAc also suggested that they used more meaningful words (based on MRC Database). As noted by McNamara et al. (2014, p.75), "words with higher meaningfulness scores are highly associated with other words (e.g., "people" (612)), whereas a low meaningfulness score (e.g., "abbess" (218)) indicates that the word is weakly associated with other words".

5.2 Logistic Regression Analysis

For further analysis and in response to the second research question, a binary logistic regression analysis was conducted in SPSS to explore the predictive ability of the lexical indices as predictor variables. It was preferred to discriminant analysis, as logistic regression is more flexible in terms of assumptions and less sensitive to the normality of the groups. Pallant (2020) explains that binary logistic regression can be used "to assess how well your set of predictor variables predicts or explains your categorical dependent variable" (p. 169), and in terms of the assumptions for the analysis, correlations among the predictor variables (multicollinearity) and outliers must be checked.

First, as shown in Table 3, the full model with all predictors was statistically significant, x2 (16, N = 217) = 169.191, p<.0005, indicating that a distinction can be made between NES vs. NNES writers in terms of all 17 lexical indices of the study.

		Chi-square	df	Sig.
Step 1	Step	169.191	16	.000
	Block	169.191	16	.000
	Model	**169.191**	**16**	**.000**

Table 3. Omnibus Tests of Model Coefficients for Lexical Indices

The model summary table (Table 4) shows that the model, as a whole, explained between 53.8% (Cox and Snell square) and 71.8% (Negelkerke R squared) variance in the writers' status, with 83.6% correct case classification as indicated in Table 5.

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	134.037a	.538	.718
a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.			

Table 4: Model Summary

Table 5 also indicates sensitivity and specificity of the model, percentage of NES writers accurately identified by the model (98%) and the percentage of NNES writers correctly classified as non-native writers (85%) respectively. Pallant (2020) refers to positive and negative predictive values as better indicators of sensitivity and specificity. In this study, positive predictive value was calculated as (98/20+98) 83% and negative predictive value was (85/85+16) 84%.

Observed			Predicted		
			Grouping		Percentage Correct
			Non-native	Native	
Step 1	Grouping	Non-native	85	20	81.0
		Native	16	98	86.0
	Overall Percentage				83.6
a. The cut value is .500					

Table 5. Classification Table

However, as indicated in the table for 'Variables in the Equation' (see Appendix), only four lexical variables made a statistically significant unique contribution to the model (WRDFRQmc, WRDFAMc, WRDHYPn and LDMTLD).

5.3 Integrating Qualitative and Quantitative Evidence of Lexical Differences

Quantitative results indicated a high demand for both academic writing proficiency and expertise among NES and NNES professional architects. However, statistically significant differences were observed, favoring the hypothetical expectation of performance differences between the two groups. Although these statistical findings are based on a large dataset, the brief qualitative analysis of the following excerpts from both groups can provide additional insights into the nature of these differences.

The first excerpt, titled "The Lighthouse 65" was released by the NES professional architects of a British firm (AR Design Studio in 2025).

> The Lighthouse is a beachfront property on the south coast of England. It is a super insulated, luxury 3 bedroom house sitting in a beautiful water side location enjoying stunning views of the Solent and the Isle of Wight. The site is enclosed between two neighboring buildings and a 7m high embankment to the north; of which pavement and street access sits at the top. This access, to the roof level of the property, and the one-directional view over the beach, English Channel and Isle of Wight to the south, has led to an interesting design which takes inspiration from traditional beach pavilions such as the De La Warr in Bexhill a little further along the coast. The design concept was to maximize the building's width, so every key room enjoys expansive views of the vista. All bathrooms and utility spaces run at the rear of the property, allowing the view to be continuous for all living spaces. The house sits 7 meters below road level with the roof acting as a parking deck for 3 cars. Visually the roof and floor decks are hung from the central concrete core, terminating in large cantilevers that provide shade and open-air shelter to the ground floor. Balconies and outdoor accommodation are provided by horizontal planes carefully cantilevered from a central access tower which penetrates the roof plane; atop of which sits a frameless glass enclosure providing access and acting as a light-house. Lighting within this enclosure is triggered by a barometer providing instantaneous information to the beach and sea beyond. This illuminated glass cube sits on top of the concrete stair core. At night the lit glass glows to indicate local weather conditions: green when weather is fair and red when atmospheric pressure drops, warning passing yachtsmen of possible stormy conditions.

The second text, "Darya [sea] Residential Building" is provided by the NNES professional architects of an Iranian firm (Hooba Design, 2025).

> Darya Building is one of the millions of Infield buildings located in Tehran, which must adhere to top-down regulations for its design. In recent years, Tehran, like many other cities in Iran, has experienced a surge in urban construction, primarily focusing on medium-scale buildings within the urban fabric. Apartment buildings have become clustered among each other, often only receiving light from one side. In this project, the façade is designed as a green and light-enhancing surface, achieved by creating a double-skinned façade. An inner layer, made from glazed materials related to the void spaces, reflects copper-colored light into the interior spaces, transforming the project's façade into a green wall. In this project, the idea of green spaces, windows, and interior spaces are integrated to create a unified structure aimed at improving both indoor and outdoor spaces. The project's spatial formation draws from two main concepts: "void" and "mass," each combined with the project's materiality. The ultimate goal is to create a replicable pattern for this type of residential apartment that can contribute to enhancing the quality of life in urban apartments and endow the project's spatial elements with a sense of sustainability. The structure created on the exterior façade forms a cohesive whole through elements with a subtle form to maintain clarity in interpreting the project. The project's materiality was formed based on two elements: Glazed brick for the void section, enhancing light reflection, and providing a unique lighting experience within the voids. Recycled stone with a grainy texture representing the project's residential volumes. This organizational approach can maximize air circulation and light within the apartments, allowing semi-open spaces to play a more significant role in the daily life of urban apartment buildings.

First, an initial review of both texts reveals that, as expected in descriptive genres, they are rich in content words, which are *italicized* for this purpose. The following table displays the frequency and percentage of content words in each text. The extensive use of nouns and adjectives, combined with high levels of lexical density (the proportion of content words to total words), is characteristic of the descriptive academic genre and reflects advanced writing style (Biber & Gray, 2010; Nasseri & Thompson, 2021).

	NES text	NNES text
Nouns	94 (33.2%)	95 (34%)
Adjectives	30 (10.6%)	32 (11.4%)
verbs	32 (11.3%)	44 (15.7%)
Adverbs	6 (2.1%)	4 (1.4%)
Function words	121 (42.8%)	105 (37.5%)
Total words	283 (100%)	280 (100%)
Lexical density	.57	.62

Table 6. Frequency and Percentage of Content Words

In terms of lexical diversity, the quantitative results build on the qualitative observations above, providing a more precise numerical perspective on the patterns identified in the excerpts. They indicated higher type–token ratios for NNES authors, with the exception of the TTR for content words. Although the example excerpts are not representative of the entire corpus, they suggest that both descriptive texts rely on nouns and adjectives as their main content words. A closer look showed, as expected, the most frequent words in both texts are function words; however, using Reuneker's (2017) Lexical Diversity Measurement tool indicated that the NES author used a higher number of function words (n = 76), including the (29), and (12), of (11), a (10), to (9), and is (5). By contrast, the NNES author's function words were fewer (n = 63), including the (21), a (12), in (9), and (7), to (7), and of (7). Differences in the proportion of verbs and function words may contribute to the observed variation in lexical diversity indices (Table 2), reflecting either the NNES authors' stylistic tendencies or an influence of L1 transfer.

Lexical sophistication involves the use of less frequent academic and technical words and phrases that are often specific to ESP texts and can indicate proficiency in specialized academic writing (Kyle & Crossley, 2016). Quantitative results, however, indicate that among the lexical-sophistication indices with statistically significant differences, three parameters showed potential to discriminate NNES and NES authors, including content-word familiarity (WRDFAMc) and content-word frequency (WRDFRQc). In the NES sample writing, words such as *property, insulated, location, access, design, and concept* are familiar and frequent, suggesting that the text can be accessible to a broader audience—non-experts attracted to visiting ArchDaily. Another lexical sophistication variable with a statistically significant contribution to the model is WRDHYPn, indicating that NES authors used more specific nouns. The technical term '*cantilever*', which refers to a specialized structural element within architectural and structural engineering contexts and appears in the NES sample, serves as a good example. Although the model did not assign a predictability value for WRDCNCc, WRDIMGc, and WRDMEAc, NNES authors significantly used more concrete, meaningful, and imageable content words such as *building, design, urban, construction, façade, and copper-colored*.

6. Discussion

The main objective of the present study was to investigate the lexical differences between NES and NNES professional architects' descriptive texts on ArchDaily. The focus of the study was on lexical diversity and sophistication indices, which proved to be theoretically and operationally relevant to L1 and L2 writing discriminations (Crossley & Kim, 2022).

The first research question examined whether the descriptive texts produced by NES and NNES professional architects differ with respect to lexical indices. In line with prior work (Crossley & Kim, 2022), preliminary results from independent-samples t-tests—a common tool in similar studies (McNamara et al., 2014)—supported the existence of lexical differences between NES and NNES architects. Specifically, the t-test results for lexical diversity indices (LDTTRa, LDMTLDa, and LDVOCDa) showed that NNES writers (with higher mean values) employed a greater variety of

unique words relative to the total number of words, indicating a higher lexical diversity in their texts. Although prior EGAP research (Crossley & McNamara, 2011; Kyle et al., 2021) and several ESAP studies (Crossley & McNamara, 2009; Jung et al., 2019; Azadnia, 2022; Crossley & Kim, 2022) often associate higher lexical diversity with greater linguistic proficiency, this study found higher lexical diversity among NNES writers. This unexpected reversal challenges conventional assumptions that lexical variety straightforwardly indexes proficiency, indicating instead that diversity may also reflect genre- or audience-related choices rather than linguistic limitation.

To explain this discrepancy, it is important to note that NNES and NES writings did not show a significant difference in LDTTRc, when the analysis is restricted to content words (nouns, verbs, adjectives, and adverbs). This pattern points to the possibility that the overall disparity may be driven more by function words, rather than by core content terms. The qualitative analysis of the sample texts corroborates the quantitative findings that lexical diversity across all words does not fully capture how word classes contribute to variation in our data. Specifically, while the full-word analysis indicates a difference between NES and NNES authors, the lexical diversity of content words—particularly nouns and adjectives—shows no significant between-group difference. This concordance between qualitative impressions and content-word metrics suggests that core descriptive vocabulary (the substantive nouns and adjectives) is remarkably similar across NES and NNES writers, even as broader, non-content word usage drives the observed divergence when all words are considered. Consequently, interpretations of NES–NNES differences in lexical diversity should be made cautiously, with attention to the word classes driving the effect. The apparent similarity in core vocabulary implies that the disparity detected at the full-word level may reflect variations in function words, discourse markers, or domain-specific terminology that do not alter the substantive content conveyed by the texts. In addition, a higher proportion of function words may enhance cohesion, thereby improving readability and comprehension (McNamara, Louwerse, McCarthy, & Graesser, 2010). This difference therefore appears to be rhetorical rather than epistemological, reflecting variation in linguistic habitus and textual convention rather than in disciplinary knowledge.

Similar to lexical diversity indices, the overarching idea, particularly in the EGAP setting, is that more proficient writers possess broader lexical knowledge and use more sophisticated, less frequent words (Chen & Baker, 2010; Lenko-Szymanska, 2014; Huang, 2015; Crossley & Kyle, 2018; Garner & Kyle, 2020). The results, however, suggest that NES architects tend to use more frequent and familiar words, while NNES writers appear to adhere to EGAP/ESAP norms to improve the quality of their descriptive texts. They achieve this by incorporating more precise and professional vocabulary and by maintaining a tone that is exact, detail-oriented, and tailored to an audience well versed in the field. By contrast, the NES style reflects accessibility-oriented choices designed for a broader readership, as non-experts increasingly visit ArchDaily. This pattern aligns with enculturation theory, which posits that NES outputs skew toward readily accessible lexicon, whereas NNES outputs exhibit enhanced lexical sophistication and vivid imagery when aligned with target disciplinary conventions (Lei & Yang, 2020; Nie, 2024). This contrast highlights that linguistic proficiency in ESP writing should not be equated with lexical sophistication alone but with rhetorical appropriateness to the communicative situation. In this sense, while NNES writers demonstrate academic refinement, NES professionals display genre awareness and sensitivity to audience expectations—an equally advanced form of expertise in disciplinary communication.

The second research question asks what measures of lexical analysis have the potential to discriminate between texts written by NES and NNES professional architects. The results addressing the question show that a small set of lexical indices reliably discriminates texts produced by NES and NNES professional architects. In the logistic regression framework, four predictors emerged as significant: WRDFRQmc, WRDFAMc, WRDHYPn, and LDMTLD. The pattern indicates that, holding other factors constant, NES and NNES writings diverge in word frequency, familiarity and specificity. Specifically, NES texts tend to exhibit higher use of frequent and familiar content words

and a broader distribution of discourse- and function-related terms. Interpreting these findings through a discourse- and genre-informed lens, NES writers appear to prioritize accessibility and cross-reader reach, aligning with ESAP-like expectations for broader audiences and general professional discourse. Collectively, the findings imply that NES writers leverage a broader, more accessible lexical profile, whereas NNES writers exhibit constructs that emphasize concreteness, and imageability. These patterns have implications for how audience, genre expectations, and professional norms shape language use in architectural discourse; they also suggest targeted directions for writing training and assessment in diverse professional populations.

Beyond diversity and sophistication, other lexical indices also differentiated NES and NNES writers. In particular, higher values for concreteness, imageability, meaningfulness, and hypernymy were found in the lexical items used by NNES writers, pointing not to linguistic limitation but to a distinct lexical strategy oriented toward precision, visualization, and disciplinary clarity. As McNamara et al. (2014) state, word concreteness refers to here-and-now concepts and objects, so their higher values in descriptive texts of the study can be a merit for NNES writers' texts. On the other hand, McNamara et al.'s norms for word concreteness from the random selection of a large corpus of scientific texts (created by the Touchstone Applied Science Associates (TASA), Inc.) indicate that word concreteness measures decrease with age and proficiency. The same trend was observed by Graesser, McNamara, and Kulikowich (2011) that science texts showed a decrease in lexical concreteness over grade levels. Imageability or the ease of constructing mental images of the word, also indicated similar results, as McNamara et al.'s norms showed the smallest values for the highest grades. Similarly, Crossley (2020) found that more proficient writers used less imageable, less familiar, less frequent, and less diverse lexical items. Similar to imageability index and in line with the results of the study, McNamara et al.'s norms for word meaningfulness showed the lowest score for the highest grade, though the results of L2 studies suggest that word meaningfulness is a significant predictor of early verb and noun production in L2 learners (Crossley & Salsbury, 2010). Words that receive higher meaningfulness scores are closely connected with other words, while a low meaningfulness score suggests that the word has weaker associations with other words (McNamara et al., 2014). It can be inferred that NNES writers were attentively trying to have suitable lexical choices. These findings indicate that NNES writers' tendency toward concreteness and imageability does not reflect limited proficiency, but rather a deliberate lexical strategy to communicate with precision and disciplinary depth.

The hypernymy indices showed that NES writers had a significantly higher mean when nouns were analyzed, while NNES writers had a higher mean when verbs were examined. According to McNamara et al. (2014), a lower hypernymy value indicates a tendency to use less-specific words, while a higher value indicates a tendency to use more-specific words. As stated by Crossley, Salsbury, McNamara, and Jarvis (2011), L2 learners acquire hypernymic relations during their cognitive developmental stages and that their ability to use specific words to convey nuanced meanings is indicative of advanced proficiency levels. Conversely, NES writers' preference for broader, more familiar nouns aligns with an accessibility-oriented strategy aimed at reaching a wider professional and non-expert audience.

Overall, the observed differences between NES and NNES writings suggest that the primary driver of lexical behavior in ESAP/ESP settings is genre-specific conventions and enculturation, underscoring that performance is governed by enculturation into the target discourse community rather than nativeness alone (Lei & Yang, 2020; Nie, 2024). Taken together, these findings highlight that expertise and disciplinary alignment, rather than L1 background, play a decisive role in shaping lexical choices across professional architectural discourse (McKinley & Rose, 2018; Lei & Yang, 2020).

7. Conclusions

Viewed through a wider lens, this study reconsiders the long-standing assumption in academic writing research that NES/NNES distinctions correspond to proficiency differences, revealing instead a more intricate relationship between lexical choice, genre expertise, and disciplinary enculturation. Rather, the results suggest that in ESP contexts, higher proficiency entails greater sensitivity to sociocultural nuance and audience comprehension and processing ease.

The implications of the study extend to professionals with high language proficiency seeking integration into their community of practice. The findings suggest that, unlike in EGAP contexts, lexical differences in ESP settings do not signify gaps in proficiency. Instead, they reflect distinct rhetorical strategies that are equally appropriate for specialized professional audiences and contexts. These findings have significant pedagogical implications, highlighting that ESP courses should prioritize the importance of tailored instruction—particularly focusing on genre-specific conventions and contextually appropriate lexical choices.

In preference to relying solely on generalized native-speaker norms, a curriculum that addresses the particular communicative practices and discourse patterns relevant to professional and specialized contexts would better prepare learners for real-world application. This approach recognizes the diversity of rhetorical strategies and vocabulary suited to different audiences and purposes, thereby promoting more effective and authentic professional communication. In other words, effective writing in context-specific genres requires adherence to the stylistic norms of the target community - an aspect that deserves careful attention in ESP settings and academic writing courses.

Another important point to consider is that NNES experts usually hire language experts to edit, translate or write their texts. If the texts labeled as NNES writings in this study are written or edited by such language professionals, then the observed discrepancy between the writings may originate from their lack of expertise in architecture and the tendency to elaborate the writing quality through overusing the context-independent and widely accepted standard criteria of advanced writing, rather than adhering to the stylistic norms of the target community of practice. Regardless of whether the texts were written by NNES architects or language experts, the results provide an important message that appropriate language use in context-dependent genres requires attention to the target audience and stylistic conventions of the relevant professional community.

The study is subject to certain limitations. The assumption that NES English architects serve as the standard benchmark was based on selecting the leading successful architecture firms as the target, while no information was available regarding the success of the NNES group being compared, nor did the study indicate whether they had hired language editors or translators to write the texts, though it is very common in the context of the study.

A related consideration concerns authorship practices among NNES professionals. NNES experts often hire language specialists to edit, translate, or draft their texts. If the texts labeled as NNES writings in this study were produced or revised by such professionals, the observed discrepancies may stem from their lack of disciplinary expertise and from a tendency to enhance linguistic quality through overreliance on general standards of advanced writing, rather than adhering to the stylistic norms of the architectural community of practice. Regardless of authorship origin, the results underscore that appropriate language use in context-dependent genres requires close attention to both audience expectations and disciplinary conventions. At the same time, even if this were the case and such texts were to be configured as mediated productions, it would not invalidate the study's findings. The differences from the native corpus should still be interpreted as outcomes of intercultural mediation—intercultural artefacts embodying the negotiation between linguistic expertise, often provided by L1 editors, and disciplinary knowledge contributed by NNES professionals that characterizes professional ESP communication.

Additionally, although this study is primarily computational and quantitative, we chose two sample texts as the unit of analysis, with the aim of retaining text-level coherence rather than focusing on sentence-level units. While this design affords some qualitative interpretation, the limited sample size restricts the robustness of qualitative inferences. Therefore, future research that incorporates more detailed qualitative or mixed-method approaches—such as combining computational lexical tools (e.g., Coh-Metrix) with qualitative genre or discourse analyses—would be valuable in providing a more comprehensive understanding of the pragmatic, rhetorical, and syntactic dimensions of professional ESP communication. Finally, the study solely focused on lexical indices, whereas a comprehensive understanding of the differences may necessitate a broader analysis encompassing various textual characteristics.

Acknowledgment and Funding

I am indebted to James Haner, Senior Research Software Developer at University of Memphis Institute of Intelligent Systems, for providing access to Cohmetrix Core Desktop Beta (2023) package that was crucial for carrying out this research. His invaluable assistance has been greatly appreciated.

I would like to acknowledge the financial support provided by Shiraz University of Arts through grant number (2002/1402) for funding this research project.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

AR Design Studio. (2025, July 6). The Lighthouse 65. Retrieved from ArchDaily: https://www.archdaily.com/363694/the-lighthouse-65-ar-design-studio?ad_source=search&ad_medium=projects_tab

ArchDaily. (2024). Retrieved from ArchDaily: https://www.ArchDaily.com/content/about?ad_source=jv-header&ad_name=hamburger_menu

Anderson, R. (2014). A parallel approach to ESAP teaching. Procedia-Social and Behavioral Sciences, 136, 194-202. https://doi.org/10.1016/j.sbspro.2014.05.313

Azadnia, M. (2021). A corpus-based analysis of lexical richness in EAP texts written by Iranian TEFL students. Teaching English as a Second Language Quarterly, 40(4), 61-90. https://doi.org/10.22099/JTLS.2021.40043.2960

Carkin, S. (2005). English for academic purposes. In Handbook of research in second language teaching and learning (pp. 85-98). Routledge.

Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. Language Learning and Technology, 14, 30-49.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian knot: The moving-average type-token ratio (MATTR). Journal of Quantitative Linguistics, 17(2), 94-100. https://doi.org/10.1080/09296171003643098

Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. Journal of Writing Research, 11(3), 415-443. https://doi.org/10.17239/jowr-2020.11.03.01

Crossley, S. A., & Kim, M. (2022). Linguistic features of writing quality and development: A longitudinal approach. Writing Analytics, 6, 59-93. https://doi.org/10.37514/JWA-J.2022.6.1.04

Crossley, S. A., & Kyle, K. (2018). Assessing writing with the tool for the automatic analysis of lexical sophistication (TAALES). Assessing Writing, 38, 46-50. https://doi.org/10.1016/j.asw.2018.06.004

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. Journal of Second Language Writing, 18, 119-135. https://doi.org/10.1016/j.jslw.2009.02.002

Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of the first and second language writing. International Journal of Continuing Engineering Education and Lifelong Learning, 21(2-3), 170-191. https://doi.org/10.1504/IJCEELL.2011.040197

Crossley, S. A., & Salsbury, T. (2010). Using lexical indices to predict produced and not produced words in second language learners. The Mental Lexicon, 5(1), 115-147. https://doi.org/10.1075/ml.5.1.05cro

Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011). Predicting lexical proficiency in language learner texts

using computational indices. Language Testing, 28(4), 561-580. https://doi.org/10.1177/0265532210378031

Crossley, S. A., Skalicky, S., Kyle, K., & Monteiro, K. (2019). Absolute frequency effects in second language lexical acquisition. Studies in Second Language Acquisition, 41(4), 721-744. https://doi.org/10.1017/S0272263118000268

Doró, K., & Pietilä, P. (2015). Researching Vocabulary in L2 Writing: Methodological Issues and. In P. Pietilä, K. Doró, & R. Pípalová, Lexical Issues in L2 Writing (pp. 11-29). Cambridge Scholars Publishing.

Dowell, N. M., Graesser, A. C., & Cai, Z. (2016). Language and discourse analysis with Coh-Metrix: Applications from educational material to learning environments at scale. Journal of Learning Analytics, 3(3), 72-95.' https://doi.org/10.18608/jla.2016.33.5

Garner, J., Crossley, S., & Kyle, K. (2020). Beginning and intermediate L2 writers' use of N-grams: An association measures study. International Review of Applied Linguistics in Language Teaching, 58(1), 51-57. https://doi.org/10.1515/iral-2017-0089

Graesser, A. C., McNamara, D. S., & Kulikowich, J. (2011). Coh-Metrix: providing multi-level analyses of text characteristics. Educational Researcher, 40(5), 223-234. https://doi.org/10.3102/0013189X11413260

Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. System, 53, 13-23. https://doi.org/10.1016/j.system.2015.06.011

Hyland, K. (2016). Academic publishing and the myth of linguistic injustice. Journal of Second Language Writing, 31, 58-69. https://doi.org/10.1016/j.jslw.2016.01.005

Jung, Y., Crossley, S., & McNamara, D. (2019). Predicting second language writing proficiency in learner texts using computational tools. Journal of Asia TEFL, 16(1), 37-52. https://doi.org/10.18823/asiatefl.2019.16.1.3.37

Kaivanpanah, S., Alavi, S. M., Bruce, I., & Hejazi, S. Y. (2021). EAP in the expanding circle: Exploring the knowledge base, practices, and challenges of Iranian EAP practitioners. Journal of English for Academic Purposes, 50, 100971. https://doi.org/10.1016/j.jeap.2021.100971

Kyle, K., & Crossley, S. (2015). 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. (2016). The relationship between lexical sophistication and independent and source-based writing. Journal of Second Language Writing, 34, 12-24. https://doi.org/10.1016/j.jslw.2016.10.003

Kyle, K., Crossley, S. A., & Jarvis, S. (2021). Assessing the validity of lexical diversity indices using direct judgement. Language Assessment Quarterly, 18(2), 154-170. https://doi.org/10.1080/15434303.2020.1844205

Lei, S., & Yang, R. (2020). Lexical richness in research articles: Corpus-based comparative study among advanced Chinese learners of English, English native beginner students and experts. Journal of English for Academic Purposes, 47, 100849. https://doi.org/10.1016/j.jeap.2020.100894

Lenko-Szymanska, A. (2014). The acquisition of formulaic language by EFL learners: A cross sectional and cross-linguistic perspective. International Journal of Corpus Linguistics, 19, 225-251. https://doi.org/10.1075/ijcl.19.2.04len

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

McKinley, J., & Rose, H. (2018). Conceptualizations of language errors, standards, norms and nativeness in English for research publication purposes: An analysis of journal submission guidelines. Journal of Second Language Writing, 42, 1-11. https://doi.org/10.1016/j.jslw.2018.07.003

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). Automated Evaluation of Text and Discourse with Coh-Metrix. Cambridge University Press. https://doi.org/10.1017/CBO9780511894664

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. Discourse Processes, 47(4), 292-330. https://doi.org/10.1080/01638530902959943

Nasseri, M., & Thompson, P. (2021). Lexical density and diversity in dissertation abstracts: Revisiting English L1 vs. L2 text differences. Assessing Writing, 47, 100511. https://doi.org/10.1016/j.asw.2020.100511

Nie, M. (2024). A Study on the Lexical Complexity of English for Special Purpose Based on WordSmith Tools and Software Range. In the proceedings of the 2024 4th International Conference on Internet Technology and Educational Technology Educational Informatization (ITEI 2024) (pp. 73-80). Atlantis Highlights in Social Sciences, Education and Humanities. Retrieved from https://doi.org/10.2991/978-94-6463-560-7_10

Pallant, J. (2020). SPSS Survival Manual: A Step-by-Step Guide to Data Analysis Using IBM SPSS. McGraw-Hill. https://doi.org/10.4324/9781003117445

Pérez-Llantada, C. (2012). Scientific Discourse and the Rhetoric of Globalization. Continuum.

Pípalová, R. (2015). Reporting Verbs in Native and Non-Native Academic Discourse. In P. Pietilä, K. Doró, & R. Pípalová, Lexical Issues in L2 Writing (pp. 127-155). Cambridge Scholars Publishing.

Read, J. (2000). Assessing Vocabulary. Cambridge University Press. https://doi.org/10.1017/CBO9780511732942

Reuneker, A. (2017). Lexical Diversity Measurements. Retrieved 10 October, 2025,from https://www.reuneker.nl/files/ld.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. TESOL Quarterly, 24(7), 657-677. https://doi.org/10.2307/3587400

Appendix

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	WRDFRQc	5.575	4.739	1.384	1	.239	263.787	.024	2850266.38
	WRDFRQa	-2.071	5.068	.167	1	.683	.126	.000	2594.790
	WRDFRQmc	-1.400	.597	5.493	1	.019	.247	.076	.795
	WRDAOAc	—.036	.019	3.470	1	.063	.964	.928	1.002
	WRDFAMc	-.175	.064	7.559	1	.006	.839	.741	.951
	WRDCNCc	.035	.031	1.255	1	.263	1.036	.974	1.101
	WRDIMGc	.069	.039	3.147	1	.076	1.072	.993	1.157
	WRDMEAc	-.018	.024	.606	1	.436	.982	.937	1.028
	WRDPOLc	1.442	.894	2.605	1	.106	4.231	.734	24.380
	WRDHYPn	-3.456	1.094	9.976	1	.002	.032	.004	.269
	WRDHYPv	1.105	1.551	.508	1	.476	3.019	.144	63.098
	WRDHYPnv	-1.276	2.029	.395	1	.530	.279	.005	14.897
	LDTTRc	4.964	9.735	.260	1	.610	143.156	.000	2766015263
	LDTTRa	-11.6	7.423	2.459	1	.117	.000	.000	18.313
	LDMTLD	.081	.024	11.12	1	.001	1.084	1.03	1.137
	LDVOCD	-.022	.025	.786	1	.375	.978	.930	1.028
	Constant	**87.23**	**41.74**	**4.368**	**1**	**.037**	**7.717E**		

According to the table, the odd ratios (OR) of WRDFREQmc, WRDFAMc, WRDHYPn and LDMTLD are .247,.839,.032, and 1.84, respectively. The ratios suggest that for a one-unit increase in WRDFREQmc, the odds of the outcome decrease and are approximately .247 times lower for each unit increase in the predictive variable, meaning that the predictor variable is associated with lower odds of the outcome occurring. For WRDFAMc, the odds ratio is 0.839, suggesting that for a one-unit increase in the predictor variable, the odds of the outcome decrease by approximately 16.1%. In other words, the predictor variable is associated with lower odds of the outcome occurring. An odds ratio of 0.032 for WRDHYPn indicates a substantial decrease (31.25 times) in the odds of the outcome for a one-unit increase in the predictor variable. And an odds ratio of 1.84 for LDMTLD means that for a one-unit increase in the predictor variable, the odds of the outcome increase by approximately 84%, and the predictor variable is associated with higher odds of the outcome occurring.