Scientific Research Publishing

Open Journal of Modern Linguistics, 2025, 15(4), 685-706
https://www.scirp.org/journal/ojml
ISSN Online: 2164-2834
ISSN Print: 2164-2818

Multidimensional Lexical Sophistication in EFL Engineering Students' Project Proposals: Implications for Genre-Based ESP Instruction

Jia Bi, Tingting Meng

Foreign Studies College, Northeastern University, Shenyang, China
Email: bij@swc.neu.edu.cn, mengtt@swc.neu.edu.cn

How to cite this paper: Bi, J., & Meng, T. (2025). Multidimensional Lexical Sophistication in EFL Engineering Students' Project Proposals: Implications for Genre-Based ESP Instruction. *Open Journal of Modern Linguistics, 15*, 685-706.
https://doi.org/10.4236/ojml.2025.154040

Received: July 25, 2025
Accepted: August 25, 2025
Published: August 28, 2025

Abstract

Most studies on the relationship between lexical sophistication and writing quality operationalize lexical sophistication as distributional property of words and focus on argumentative essays. This study examines the various aspects of lexical sophistication of EFL engineering students' project proposals, comparing them with argumentative essays to explore genre-specific lexical features. Twenty-two indices of various dimensions of lexical sophistication were extracted from both project proposals and argumentative essays written by students with comparable English proficiency. Independent samples *t* tests revealed significant differences in lexical sophistication between the two genres. An exploratory factor analysis grouped the 22 indices of project proposals into five components. A stepwise multiple regression analysis using these components as predictors suggested that lexical sophistication exhibited limited explanatory power for project proposal quality, contrasting with its established role in general L2 writing contexts. The findings suggest pedagogical implications for ESP instruction, highlighting the integration of discipline-specific terminology, structural coherence, and contextualized function word usage to enhance engineering students' genre-specific communicative efficacy in project proposal writing.

Keywords

Lexical Sophistication, Genre, Project Proposal, Writing Quality, ESP Instruction

1. Introduction

The relationship between linguistic features and EFL learners' writing quality has been of great interest to scholars in recent years. Among all the examined linguistic features, lexical richness is an important indicator for the quality of written

productions, which includes lexical diversity, lexical density, and lexical sophistication (Lei & Yang, 2020; Zhang et al., 2021). Lexical diversity reflects the variety of words used in a text; it is usually measured by the type-token ratio. Lexical density measures the density of information in a text as it calculates the proportion of content words in a text. Different from the other two lexical features, lexical sophistication is the most controversial one, as there is no consistent definition and operationalization for it in existing literature. Most studies have operationalized it as less frequent words based on word frequency bands derived from large corpora, but recently, an increasing number of researchers have begun to examine it as a multidimensional construct, incorporating various other features besides frequency.

As an assumed factor contributing to the quality of EFL writing, lexical sophistication has been investigated by many scholars (e.g., Higginbotham & Reid, 2019; Kyle & Crossley, 2016; Kim et al., 2018; Vögelin et al., 2019) and they have confirmed that some aspects of lexical sophistication could explain a proportion of the variance in the holistic scores of written texts. However, much of the existing research has focused on common academic genres, such as argumentative, narrative, and expository essays, and fewer studies have explored lexical sophistication in discipline-specific genres. Nesi and Gardner (2012) identified 13 genre families in university-level assessed writing, among which *Proposal* is a genre that prepares students for professional practice. Although some studies have examined lexical features in grant proposals (e.g., Connor & Mauranen, 1999; Flowerdew, 2016), research on project proposals remains scarce. Project proposals, characterized by their persuasive and technical nature, require precise lexical choices to effectively communicate problem-solving strategies and solutions (Nesi & Gardner, 2012). Given these distinctive communicative demands, it is important to examine whether and how lexical sophistication operates differently in engineering students' project proposals compared to more general academic writing.

To address this research gap, the current study adopts a multidimensional and quantitative approach to investigate lexical sophistication by comparing engineering students' project proposals with argumentative essays. Specifically, it aims to identify differences in lexical patterns between the two genres and to explore whether lexical sophistication can predict proposal quality. Clarifying these issues can provide valuable insights for developing more targeted vocabulary instruction in ESP contexts, particularly within genre-based pedagogical frameworks.

2. Literature Review

2.1. Dimensions of Lexical Sophistication

Lexical sophistication is a complex construct without a universally accepted definition. Read (2000) described it as "a measure of the proportion of relatively unusual or advanced words in the learner's text" (p. 203). Due to the subjective nature of defining "unusual or advanced words," researchers have adopted different operationalizations. Early studies, such as Laufer and Nation (1995), measured

lexical sophistication as the proportion of low-frequency words in a text. Their Lexical Frequency Profile (LFP) categorized words based on frequency bands, distinguishing between the 2,000 most common words and "Beyond 2000" words, which include academic and off-list words. This classification has since been widely used to assess lexical sophistication (e.g., Zheng, 2016; Lei & Yang, 2020; Tracy-Ventura, 2017; Higginbotham & Reid, 2019; Jarvis, 2013; Juanggo, 2018). Although word frequency remains a widely used criterion, some scholars (e.g., Bulté & Housen, 2012; Kyle et al., 2018) argue for a broader perspective that considers lexical learning difficulty. More recent studies (e.g., Kyle & Crossley, 2015; Kim et al., 2018; Lu & Hu, 2021; Sun & Lu, 2021; Baese-Berk et al., 2021; Zhou et al., 2023) conceptualize lexical sophistication as a multidimensional construct, incorporating distributional, formal, psycholinguistic, acquisitional, semantic, and n-gram properties.

Distributional properties. Distributional properties include word frequency, range, contextual distinctiveness, and the proportion of academic words and phrases (Lu & Hu, 2021). Words in a text can be categorized by frequency bands derived from large reference corpora, with lower-frequency words generally considered more sophisticated. Range refers to the number of different texts in which a word occurs within a corpus; a wider range suggests a lower level of lexical sophistication. Contextual distinctiveness measures how broadly a word is used across different contexts (McDonald & Shillcock, 2001a). It can be quantified by the number of sub-corpora a word appears in or by word association tasks (McDonald & Shillcock, 2001b). The use of academic words (Coxhead, 2000) and academic phrases (Simpson-Vlach & Ellis, 2010) also indicates lexical sophistication, as advanced learners are more likely to incorporate these elements into their writing (Durrant et al., 2019; Lu & Hu, 2021).

Psycholinguistic properties. Psycholinguistic dimensions encompass concreteness, imageability, and familiarity (Coltheart, 1981; Salsbury et al., 2011). These characteristics influence word processing and learnability and are usually defined through behavioral studies with L1 speakers. Concreteness refers to how tangible or abstract a word is. Imageability reflects how easily a word evokes a mental image. Familiarity indicates how frequently a word is encountered by learners. Research suggests that advanced EFL learners tend to use less concrete, less imageable, and less familiar words (Kim et al., 2018; Crossley & Skalicky, 2019).

Formal properties. Formal properties, such as phonological and orthographic neighborhoods, contribute to word sophistication (Andrews, 1989; Kyle et al., 2018). Phonological neighborhood involves words differing by one phoneme (e.g., reverse vs. rehearse), while orthographic neighborhood refers to words differing by one letter (e.g., face vs. fact). Words with fewer phonological and orthographic neighbors are generally more sophisticated and more frequently used by proficient learners.

Acquisitional properties. Acquisitional properties include age of acquisition

(AoA) and age of exposure (AoE), measures that reflect when words are typically learned (Kuperman et al., 2012; Dascalu et al., 2016). Words acquired later in life tend to be more sophisticated.

Semantic properties. Semantic properties can be assessed through polysemy and hypernymy (Crossley et al., 2009; Crossley et al., 2010). Polysemy refers to the number of senses a word has. Advanced learners tend to use words with fewer senses, indicating higher sophistication (Guo et al., 2013). Hypernymy represents hierarchical relationships between superordinate words (e.g., fruit) and subordinate words (e.g., banana), with lower-position words in the hypernymy hierarchy being more advanced.

N-gram properties. N-grams, or multi-word units, are recurring sequences of words (e.g., the fact that). Researchers include n-gram properties such as frequency, range, and association strength in their studies to measure lexical sophistication (Eguchi & Kyle, 2020; Kim et al., 2018).

Taken together, these dimensions underscore the complexity of lexical sophistication and the importance of multidimensional analyses that go beyond simple frequency measures. Understanding these various aspects not only provides a richer description of lexical usage in L2 writing, but also opens up questions about how such lexical features relate to broader measures of writing performance.

2.2. Relationship between Lexical Sophistication and Quality of L2 Writing

Numerous studies have explored lexical sophistication in EFL writing, many of which rely on frequency indices, such as the Beyond-2000 ratio (Lei & Yang, 2020; Higginbotham & Reid, 2019) or word range metrics (Vögelin et al., 2019). With the advent of advanced computational tools, scholars have expanded their analyses to include a broader range of vocabulary features. For instance, Kim et al. (2018) employed principal component analysis to condense numerous multidimensional indices into 12 components, with the top seven explaining 24.6% of the variance in L2 writing proficiency. Similarly, Kyle and Crossley (2016) demonstrated that certain features, including range, bigrams, hypernymy, and imageability, are strong predictors of essay quality in TOEFL independent writing tasks, although their predictive power diminishes in integrated tasks.

Despite progress in lexical sophistication research, its focus remains largely on argumentative essays, making it difficult to generalize the results to other types of writing. Although some studies examine other genres, such as research articles (Lei & Yang, 2020) and application letters (Zhang et al., 2021), the diverse range of writing tasks required of EFL learners in higher education remains underrepresented. Among these, project proposals have received little attention, despite their importance in both academic and professional contexts. This study seeks to fill that gap by examining lexical sophistication in project proposals written by EFL engineering students, with implications for both pedagogy and future research.

2.3. Lexical Sophistication and Genre in ESP

Genre theory emphasizes that different genres require distinct lexical patterns to fulfill their communicative purposes (Swales, 1990; Hyland, 2004; Cheng, 2025). For instance, engineering project proposals often demand precise technical terminology and formulaic expressions to convey methodological rigor and persuasiveness (Nesi & Gardner, 2012), whereas argumentative essays generally rely on broader academic vocabulary and various rhetorical devices to build and support complex arguments. Some studies have examined how rhetorical structures in ESP genres reflect both communicative goals and disciplinary norms. For example, Connor and Mauranen (1999) analyzed persuasive moves in EU grant proposals, Flowerdew (2016) looked into genre-based approaches to teaching postgraduate proposal writing, and Kanoksilapatham (2015) investigated rhetorical variation across engineering subfields. Although these studies underscore the necessity of genre-sensitive instruction, they have not sufficiently addressed how rhetorical purposes translate into measurable lexical sophistication patterns, particularly in specialized genres such as engineering project proposals.

Given this research gap, it is essential to investigate how lexical sophistication manifests in project proposals, especially through multidimensional indices involving distributional, psycholinguistic, formal, acquisitional, semantic, and n-gram features. Examining these lexical dimensions in texts designed to propose technical solutions and persuade stakeholders will contribute to understanding how genre conventions influence precise vocabulary usage. Furthermore, comparing project proposals with argumentative essays can illuminate genre-specific lexical differences, thereby assisting ESP instructors in developing more effective, discipline-specific teaching strategies. Specifically, this study is guided by the following three research questions:

1) What are the differences in lexical sophistication between project proposals and argumentative essays written by EFL engineering students?

2) What latent components underlie the lexical sophistication indices in project proposals, and how do these components capture macro-lexical features?

3) To what extent do the resulting components of lexical sophistication from RQ2 predict the holistic scores of students' project proposals?

3. Methods

3.1. Learners' Corpora

IT Professional English is an ESP course for software engineering undergraduates at a Chinese university, designed to enhance workplace English proficiency in IT contexts. A core assignment requires students to draft a software project proposal that describes a sociotechnical system in which a computer system interacts with humans or the physical world to solve real-life problems and fulfill human needs. The proposal must identify a customer pain point, propose a solution with clear business values, and persuade hypothetical investors for funding. Students completed the assignments independently outside class, with permission to use refer-

ence sources and no time or length restrictions.

The corpus of project proposals consists of 103 texts totaling 95,808 tokens. To ensure the reliability of the evaluation, each proposal was independently assessed by two instructors with extensive experience in English writing pedagogy, using a holistic rubric covering three aspects: task completion, coherence and cohesion, and lexical and grammatical accuracy. Scores ranging from 0 to 10 were assigned based on the overall quality. The average of the two raters' scores was used as the final grade, unless their scores differed by more than two points, in which case a third instructor was consulted to resolve the discrepancy. Inter-rater reliability was assessed using the intraclass correlation coefficient (ICC), with a resulting value of .81, indicating good consistency between raters.

For comparison, an additional corpus of 318 argumentative essays (totaling 91,573 tokens) was compiled from students whose average College English Test Band 4 (CET-4) scores are comparable to those of the project proposals' authors, indicating similar overall English proficiency. The essays were written outside of class in response to various prompts, with each essay containing no fewer than 120 words. Both corpora were collected between 2019 and 2021, prior to the public accessibility of advanced large language models such as ChatGPT. While students could consult conventional resources, none of the texts were generated or linguistically refined by AI tools. Descriptive statistics for the two corpora are presented in Table 1.

Table 1. Descriptive statistics of the two corpora.

Genres	Number of texts	Words per text			Score distribution*			
		Max	Min	Mean	Max	Min	Mean	St. Dev
Project proposals	103	4044	122	930	9.15	4.75	7.56	.83
Argumentative essays	318	850	119	288	NA	NA	NA	NA

*The score distribution for argumentative essays was not recorded as it was not required in the present study.

3.2. Tools for Data Extraction and Analysis

The Tool for the Automatic Analysis of Lexical Sophistication (TAALES; Kyle et al., 2018) was used to extract various quantified lexical features from the texts in the two genres. The latest version of TAALES provides more than 1000 indices representing different aspects of lexical sophistication. For each lexical property, TAALES offers indices derived from various reference corpora. Researchers have explored these indices to gain a deeper understanding of lexical sophistication. Durrant et al. (2019) confirmed that frequency indices derived from different TAALES reference corpora were strongly correlated with each other. They also suggested that mean frequency and range measures combining content and function words might not be very meaningful, as they conflated different constructs. Moreover, the use of function words mainly reflects differences in syntactic structures rather than in vocabulary. Therefore, this study selected a single corpus-

based index for each dimension of lexical sophistication and examined content and function words separately. The selected indices were aligned with the six theoretical dimensions discussed in the literature review. Other TAALES indices outside these dimensions were excluded due to conceptual overlap or little relevance to the genre-based lexical analysis in this study.

Table 2 presents the indices selected for analyzing and comparing the two genres.

Table 2. Indices of lexical sophistication used in this study.

Property	Dimension	Label	Description
Distributional	Frequency	Freq_CW	Mean frequency score of content words
		Freq_FW	Mean frequency score of function words
	Range	Range_CW	Mean range score of content words
		Range_FW	Mean range score of function words
	Contextual Distinctiveness	CD_CW	Contextual distinctiveness score of content words
		CD_FW	Contextual distinctiveness score of function words
	Academic Words	AWL	Normed count of AWL words
	Academic Formulas	AFL	Normed count of AFL
Psycholinguistic	Concreteness	Concreteness_CW	Mean unigram concreteness score of content words
		Concreteness_FW	Mean unigram concreteness score of function words
	Familiarity	Familiarity_CW	Mean unigram familiarity score of content words
		Familiarity_FW	Mean unigram familiarity score of function words
	Imageability	Imageability_CW	Mean unigram imageability score of content words
		Imageability_FW	Mean unigram imageability score of function words
Formal	Orthographic Neighbors	Ortho_N_AW	Average number of orthographic neighbors
	Phonological Neighbors	Phono_N_H_AW	Average number of phonological neighbors
Acquisitional	Age of Acquisition	AoA_CW	Mean age of acquisition score of content words
		AoA_FW	Mean age of acquisition score of function words
Semantic	Polysemy	Content_poly	Average number of senses for content words
	Hypernymy	Hyper_verb_noun	Average number of superordinate terms for nouns and verbs
N-gram	Bigram	Bigram_Freq	Mean frequency score of bigrams
	Trigram	Trigram_Freq	Mean frequency score of trigrams

After extracting the values of the selected lexical sophistication indices, SPSS 23 was used to conduct *t* tests, exploratory factor analysis, and multiple regression analysis.

3.3. Calculation and Interpretation of Lexical Sophistication Index Values

To assess distributional properties, four quantifiable indices, frequency, range,

contextual distinctiveness, and academic language, were compared across the two genres. Frequency and range indices were selected based on the written British National Corpus (BNC Consortium, 2007). TAALES calculates each text's mean word frequency score by dividing the total frequency scores of all words by the number of words in the text. A higher frequency score indicates that a word appears more frequently in the reference corpus. Thus, a lower mean word frequency score for a text suggests a higher level of lexical sophistication. TAALES calculates the mean range scores in a similar way. Contextual distinctiveness (CD) was operationalized through association values derived from McDonald and Shillcock's (2001b) project. Words with lower contextual diversity have higher CD values, exhibit longer lexical decision latencies, and are therefore considered more sophisticated. For the aforementioned three indices, separate calculations were performed for content words and function words within each text. The other two dimensions of distributional properties are the normed scores for academic words and academic formulas, derived respectively from Coxhead's (2000) Academic Word List and Simpson-Vlach and Ellis's (2010) Academic Formulas List. TAALES calculates these scores by dividing the number of tokens found in the lists by the total word count in a text.

For psycholinguistic properties, indices of concreteness, familiarity, and imageability were extracted based on the MRC psycholinguistic database (Coltheart, 1981). TAALES calculates the mean concreteness score by dividing the sum of concreteness scores by the number of words in the text that have valid concreteness ratings. Familiarity and imageability scores are calculated in the same manner. For all three indices, lower scores indicate that the words in the text are less concrete, less familiar, and less imageable, and therefore suggest a higher level of lexical sophistication.

The formal properties examined include indices of orthographic and phonological neighbors. TAALES provides norms for these indices based on data from the English Lexicon Project (Balota et al., 2007). Words are analyzed collectively without distinguishing between content words and function words. A higher score indicates that a word has more orthographic or phonological neighbors, which is interpreted as a lower level of lexical sophistication.

TAALES provides age of acquisition (AoA) information for 30,121 lemmas collected by Kuperman et al. (2012) as a measure of acquisitional properties. Higher AoA scores indicate that the texts contain words that are typically acquired at a later age, suggesting a higher level of lexical sophistication.

Regarding semantic properties, TAALES calculates polysemy and hypernymy indices based on the WordNet database (Fellbaum, 1998). In this study, the polysemy index was defined as the mean number of senses for all content words, while the hypernymy index was defined as the mean number of superordinate terms for nouns and verbs combined. Higher polysemy values suggest that a text contains more words with multiple meanings. Higher hypernymy values imply that a text includes more words with a greater number of superordinate terms. Overall, texts

characterized by lower average polysemy and higher average hypernymy values are considered to exhibit greater lexical sophistication.

N-gram indices used in this study include mean frequency scores of bigrams and trigrams as provided by TAALES. These scores are calculated by dividing the sum of the frequency scores for bigrams/trigrams in a given text (based on bigram/trigram frequency lists derived by Crossley et al. (2012) from the written subsets of the BNC) by the total number of bigrams/trigrams in that text that have available frequency scores.

3.4. Statistical Analyses

To address the first research question, independent samples *t* tests were conducted to determine whether indices of lexical sophistication exhibited statistically significant differences between the two genres. Prior to conducting the *t* tests, the data were examined for normality by assessing skewness and kurtosis and inspecting histograms and Q-Q plots. To control for Type I error across the 22 independent samples *t* tests, a Bonferroni correction was applied, adjusting the significance threshold to .002 (i.e., .05/22). Only *p* values below this adjusted threshold were considered statistically significant in subsequent analyses.

For the second research question, an exploratory factor analysis (EFA) was conducted to better understand the relationships among various indices. Although the 22 variables fall into six categories, no established theories define the number of factors characterizing lexical sophistication in any genre, particularly in the less studied proposal genre. Previous studies (e.g., Kim et al., 2018; Eguchi & Kyle, 2020) have confirmed that various indices obtained from TAALES are highly correlated. Therefore, reducing these dimensions into a smaller set of factors is both necessary and theoretically justified to better capture the distinctions in writing quality. Conducting an EFA not only addresses the issue of multicollinearity but also identifies components that group related variables together for subsequent analyses.

Several key considerations must be taken into account when conducting an EFA, including the appropriateness of applying EFA, the method of factor extraction, the criteria for factor retention, the method of factor rotation, and the interpretation and labeling of factors (Williams et al., 2010; Plonsky & Gonulal, 2015; Loewen & Gonulal, 2015). The following discussion will explore each of these aspects in detail, outlining the procedures used in this study.

The suitability of the data for factor analysis was assessed using the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity. Data are considered appropriate for exploratory factor analysis if the KMO value exceeds .80 and Bartlett's Test of Sphericity is significant ($p < .05$) (Field, 2018; Williams et al., 2010).

Once the suitability was confirmed, principal components analysis (PCA) was used to extract factors, a method commonly employed in L2 research to reduce the data to a more manageable size (Loewen & Gonulal, 2015). To determine the

optimal number of factors, multiple criteria were considered, following standard practice in factor analysis (Hair et al., 2010). Specifically, this study applied Kaiser's criterion (eigenvalues > 1.0), cumulative percentage of variance explained, and the scree plot examination to guide factor retention.

Regarding the factor rotation method, this study adopted direct oblimin rotation based on theoretical considerations. Plonsky and Gonulal (2015) argued that oblique rotation is more appropriate for L2 research because it allows factors to be correlated, whereas orthogonal rotation assumes that factors are uncorrelated. As noted above, the variables in this study related to language learning and human cognition tend to be interrelated; therefore, direct oblimin rotation was employed to produce the factor loading matrix.

The third research question examined the extent to which lexical sophistication components could predict holistic scores of EFL engineering students' project proposals. A stepwise multiple regression analysis was conducted using the factor scores derived from the exploratory factor analysis in RQ2. Stepwise regression was chosen to address potential multicollinearity among intercorrelated factors and to systematically identify the subset of predictors most relevant to project proposal quality, thereby balancing model parsimony with explanatory power.

4. Results

4.1. RQ1: Lexical Sophistication across Genres

Normality tests confirmed that all lexical sophistication indices met the assumption of normal distribution across both genres, thereby justifying the use of *t* tests for statistical comparisons. As shown in **Table 3**, independent samples *t* tests comparing lexical sophistication between project proposals and argumentative essays revealed significant differences across multiple dimensions, supported by large effect sizes (Cohen's $d > .8$ in absolute value in most cases), suggesting robust genre-specific lexical patterns. Notably, Levene's tests for homogeneity of variances showed that only CD_FW, Familiarity_FW, Ortho_N_AW, AoA_CW, and Bigram_Freq met the equal variances assumption ($df = 419$). For the remaining variables, the assumption of homogeneity was violated, necessitating Welch's correction with adjusted degrees of freedom.

4.1.1. Distributional Properties

For content words, project proposals exhibited significantly lower frequency ($t = -14.290$, $p < .001$, Cohen's $d = -1.516$) and narrower range ($t = -19.041$, $p < .001$, Cohen's $d = -2.031$) compared to argumentative essays. Lower frequency and range scores indicate more sophisticated lexical items, suggesting that EFL students tend to use less frequent and more specialized words in project proposals compared to argumentative essays. Additionally, higher contextual distinctiveness ($t = 14.749$, $p < .001$, Cohen's $d = 1.548$) in project proposals implies that these words are used in more specific contexts, further supporting the notion of higher lexical sophistication in project proposals.

Table 3. Results of independent samples *t* tests.

Indices	Project proposals		Argumentative essays		*t*	*df*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*				
Freq_CW	.378	.086	.526	.108	−14.290	214	.000	−1.516
Freq_FW	25.402	3.646	17.110	2.951	20.968	148	.000	2.500
Range_CW	50.894	4.062	60.264	5.105	−19.041	215	.000	−2.031
Range_FW	97.924	.888	95.836	1.478	17.323	291	.000	1.713
CD_CW	.941	.084	.786	.114	14.749	234	.000	1.548
CD_FW	.903	.096	.908	.112	−0.471	419	.638	−0.048
AWL	.102	.038	.049	.025	13.344	133	.000	1.648
AFL	.062	.022	.067	.038	−1.458	296	.146	−0.161
Concreteness_CW	382.894	15.824	353.644	20.266	15.160	219	.000	1.609
Concreteness_FW	232.113	5.130	243.771	7.949	−17.296	270	.000	−1.743
Familiarity_CW	566.310	7.991	579.061	6.239	−14.799	144	.000	−1.779
Familiarity_FW	611.228	1.539	610.534	1.864	3.417	419	.001	0.406
Imageability_CW	403.076	15.302	382.857	17.723	11.196	198	.000	1.221
Imageability_FW	247.879	7.744	269.737	9.833	−23.218	217	.000	−2.470
Ortho_N_AW	7.005	.680	8.289	.784	−14.907	419	.000	−1.750
Phono_N_H_AW	15.762	2.006	19.707	2.426	−16.444	207	.000	−1.772
AoA_CW	7.381	.496	6.334	.527	17.772	419	.000	2.046
AoA_FW	4.455	.081	4.522	.113	−6.605	239	.000	−0.682
Content_Poly	7.720	.637	8.745	.996	−12.196	272	.000	−1.226
Hyper_verb_noun	4.744	.301	3.962	.392	21.170	223	.000	2.238
Bigram_Freq	.392	.097	.230	.093	15.149	419	.000	1.705
Trigram_Freq	.016	.004	.015	.006	1.476	278	.141	0.196

In contrast, function words in project proposals tended to be more frequent (t = 20.968, p < .001, Cohen's d = 2.500) and exhibited a broader range (t = 17.323, p < .001, Cohen's d = 1.713), suggesting that function words in this genre are less sophisticated than those in argumentative essays. This indicates that project proposals rely on common structural elements (e.g., *the*, *and*, *of*) to organize informational content, while argumentative essays use less frequent function words (e.g., *whereas, hence, furthermore*), possibly reflecting the rhetorical demands of argumentative writing, which require more specialized connectives and transitional phrases to structure arguments. Additionally, there was no significant difference in the contextual distinctiveness of function words between genres (t = −.471, p = .638, Cohen's d = −.048), implying that function words in both proposals and essays are used in similarly diverse contexts.

In terms of academic language, project proposals contained more academic words than argumentative essays (t = 13.344, p < .001, Cohen's d = 1.648), whereas the two genres did not differ significantly in their use of academic formulas (t = -1.458, p = .146, Cohen's d = −.161).

4.1.2. Psycholinguistic Properties

Regarding content words, project proposals scored higher on concreteness (t = 15.160, p < .001, Cohen's d = 1.609) and imageability (t = 11.196, p < .001, Cohen's d = 1.221), but lower on familiarity (t = −14.799, p < .001, Cohen's d = −1.779). These results indicate that project proposals prioritize precise, tangible, and less common vocabulary, whereas essays utilize more abstract and familiar terms, likely to appeal to broader audiences.

However, function words displayed the opposite trend: those in argumentative essays were more concrete and imageable, whereas those in project proposals were slightly more familiar. This suggests that the function words in essays may include a greater proportion of semantically meaningful elements that contribute to explicit argumentation and reader engagement. In contrast, the function words in project proposals may be more conventionalized and predictable, reflecting the formulaic and structured nature of this genre.

4.1.3. Formal Properties

Words in project proposals had fewer orthographic (t = −14.907, p < .001, Cohen's d = −1.750) and phonological neighbors (t = −16.444, p < .001, Cohen's d = −1.772), reflecting the use of lexically unique or specialized terms with fewer similar spellings or pronunciations. This further underscores the sophisticated nature of proposal vocabulary.

4.1.4. Acquisitional Properties

The AoA score for content words was significantly higher in project proposals (t = 17.772, p < .001, Cohen's d = 2.046), indicating a preference for advanced, late-learned words. In contrast, function words in project proposals exhibited lower AoA scores (t = −6.605, p < .001, Cohen's d = −.682) compared to those in argumentative essays. This contrast highlights divergent acquisitional patterns between content and function words across genres, with project proposals favoring more specialized, late-acquired content vocabulary while relying on relatively basic, early-acquired function words.

4.1.5. Semantic Properties

For the polysemy index, project proposals exhibited significantly fewer polysemous words compared to argumentative essays (t = −12.196, p < .001, Cohen's d = −1.226), reflecting a preference for unambiguous terminology. In terms of hypernymy, project proposals exhibited a significantly higher hypernymy index (t = 21.170, p < .001, Cohen's d = 2.238), indicating that vocabulary in project proposals is semantically more specific, with terms nested under multiple layers of superordinate categories. This hierarchical structuring of terminology systemati-

cally enhances clarity and reduces ambiguity by embedding specialized concepts within precise conceptual frameworks.

4.1.6. N-Gram Properties

Project proposals had a higher proportion of bigrams listed in the written BNC (t = 15.149, p < .001, d = 1.705), suggesting an abundance of discipline-specific collocations (e.g., research objectives, budget allocation). No significant difference was observed for trigrams, possibly due to the limited variability in longer fixed phrases across genres.

Overall, project proposals demonstrate markedly higher lexical sophistication than argumentative essays across most dimensions. This aligns with their communicative purpose: proposals aim to persuade sponsors through precise, technical, and contextually distinct language, whereas essays prioritize argumentative clarity using familiar and flexible vocabulary. The findings corroborate Nesi and Gardner's (2012) observation that proposal writing is inherently informational, often featuring numerous nouns, longer words, and a high type-token ratio, and thus requiring discipline-specific terminology rarely found in general English.

4.2. RQ2: Latent Components of Lexical Sophistication in Project Proposals

An exploratory factor analysis (EFA) with direct oblimin rotation was conducted to identify latent variables underlying the 22 lexical sophistication indices. The Kaiser-Meyer-Olkin measure (KMO = .818) and Bartlett's test of sphericity (p < .001) confirmed the suitability of the data for factor analysis. Using a combination of Kaiser's eigenvalue criterion (eigenvalues > 1.0), scree plot, and the total variance explained, five factors were retained, collectively explaining 74.65% of the total variance. A minimum factor loading threshold of .40 was used to interpret the factor structure, following standard practice (Hair et al., 2010). All retained indices exceeded this threshold and loaded saliently on a single factor, with no substantial cross-loadings observed. In addition, all variables had communalities above .40, indicating sufficient shared variance with the extracted factors to justify inclusion in the factor solution. **Table 4** displays the rotated factor loadings.

Factor 1 included 10 indices: four indices reflecting distributional properties of content words, two indices of formal properties for all words, two indices capturing semantic properties exclusive to content words, as well as age of acquisition and familiarity of content words. Although the indices for phonological and orthographic neighbors were computed for all words, most of these neighbors pertained to content words. Therefore, Factor 1 was labeled "content word properties".

Factor 2 was characterized by two distributional metrics of function words, age of acquisition of function words, familiarity of function words, and bigram frequency. These collectively operationalize functional elements in discourse structuring; therefore, this factor was labeled "function word properties".

Factor 3 consisted of two highly correlated variables, imageability and concreteness

Table 4. Rotated factor structure matrix for lexical sophistication indices of project proposals (principal component analysis with oblimin rotation).

Indices	Factors				
	1	2	3	4	5
AoA_CW	−.889				
Phono_N_H_AW	.876				
Ortho_N_AW	.870				
CD_CW	−.832				
Range_CW	.803				
AWL	−.784				
Familiarity_CW	.774				
Hyper_verb_noun	−.723				
Freq_CW	.695				
Content_poly	.635				
AoA_FW		−.898			
Familiarity_FW		.799			
Freq_FW		.793			
Range_FW		.710			
Bigram_Freq		.657			
Imageability_CW			.958		
Concreteness_CW			.934		
Trigram_Freq				.871	
AFL				.717	
Imageability_FW					.703
Concreteness_FW					.688
CD_FW					.573

of content words (r = .888), which represent psycholinguistic dimensions. It is logical that a word that is more imageable is also perceived as more concrete. This factor was labeled "content word imagery".

Factor 4 was defined by two multi-word unit variables: trigram frequency and the academic formulas list. The two variables, which are moderately correlated (r = .370), together form a meaningful construct labeled "multi-word units".

Factor 5 comprised the imageability and concreteness of function words along with their contextual distinctiveness. This factor was thus labeled "function word context".

The five-factor solution effectively consolidated the originally fragmented 22 indices into interpretable macro-lexical constructs. This dimensionality reduction

enhances analytical efficiency for subsequent regression modeling while preserving critical linguistic features essential for characterizing proposal discourse.

4.3. RQ3: Predictive Validity of Lexical Sophistication Factors for Project Proposal Quality

To examine the predictive power of lexical sophistication, a stepwise multiple regression analysis was performed using the five extracted factors from RQ2. Table 5 presents the correlations between the five lexical sophistication factors and holistic project proposal scores. Only Factor 1 (content word properties) and Factor 5 (function word context) demonstrated statistically significant, albeit weak, correlations with holistic scores ($r = -.221$ and .183, respectively; $p < .05$). As detailed in Table 6, the stepwise regression model revealed that these two factors collectively explained merely 7.7% of the variance in project proposal quality (Adjusted $R^2 = .077$), a stark contrast to Kim et al.'s (2018) finding that lexical sophistication accounted for 24.6% of variance in L2 writing proficiency.

Table 5. Correlations between factor scores and project proposal scores.

Factors	Correlation with Project proposal scores	Sig. (1-tailed)
1. Content word properties	−.221	.012
2. Function word properties	.134	.089
3. Content word imagery	.153	.062
4. Multi-word units	.139	.08
5. Function word context	.183	.032

Table 6. Results of multiple regression for lexical sophistication factors predicting project proposal quality.

Entry	Predictors included	R	Adjusted R^2	R^2 Change	*B*	*SE*	β
1	F1: Content word properties	.221	.039	.049	−.206	.079	−.250
2	F5: Function word context	.308	.077	.046	.179	.079	.216

Note. Estimated constant term = 7.557; *B* = unstandardized beta; *SE* = standard error; β = standardized beta.

The limited explanatory power of the regression model suggests that lexical sophistication indices may have reduced relevance in predicting genre-specific writing quality. Unlike general L2 writing tasks where vocabulary features explain substantial variance, project proposals likely prioritize discipline-specific rhetorical or structural criteria over lexical complexity. Consequently, generic lexical metrics may inadequately capture the multidimensional criteria governing proposal quality assessments.

5. Discussion

The results of this study reveal distinct lexical sophistication patterns between project proposals and argumentative essays, with limited predictive utility of these features for proposal quality. The discussion below elaborates on genre-driven lexical variation, clustering mechanisms of indices, and the diminished role of lexical sophistication in technical writing assessment.

5.1. Genre-Driven Lexical Variation

The independent samples *t* tests demonstrated significant differences in 19 out of 22 lexical sophistication indices between project proposals and argumentative essays. This underscores the decisive role of genre differences in lexical selection, despite comparable English proficiency levels among students.

As theorized by Nesi and Gardner (2012), genres are defined by their social functions and structural frameworks. Argumentative essays aim to develop logical reasoning through a three-stage structure (introduction, argumentation, conclusion), while project proposals emulate professional practice by adopting a problem-solution-evaluation-action plan framework, each demanding distinct lexical strategies. Biber's (1988) multidimensional analysis further clarifies these distinctions. Proposals tend to score higher than essays in terms of informational density and persuasiveness, with essays falling somewhere in between. This difference is also reflected in word choice. For example, software project proposals often include technical terms like *algorithm optimization* or *scalability* when describing solutions, and use procedural vocabulary such as *milestones* and *budget allocation* in outlining plans. In contrast, essays rely more on argumentative markers like *therefore* or *conversely*, and abstract nouns such as *justice* and *equity*. These differences highlight how genre conventions influence the way writers use language.

5.2. Clustering of Lexical Sophistication Indices

While TAALES generates over 1000 indices, redundancy arises from overlapping measurement dimensions and corpus dependencies. Previous attempts to address this issue have primarily focused on intra-category clustering. Kim et al. (2018) condensed hundreds of TAALES indices into 12 components for L2 writing proficiency prediction, with each component aggregating 3 - 14 indices measuring similar lexical features. Similarly, Eguchi and Kyle (2020) grouped 110 frequency measures into 10 factors, demonstrating that indices sharing measurement dimensions but derived from different corpora consistently clustered together. These studies established that redundancy in TAALES outputs stems primarily from corpus-specific variations of equivalent measurement constructs.

Building on these findings, this study extends the investigation to cross-category clustering patterns through factor analysis. Notably, Factor 1 combines 10 indices spanning five theoretically distinct categories: distributional, psycholinguistic, formal, semantic, and acquisitional dimensions, illustrating interdependencies among lexical properties. Words with fewer orthographic neighbors are

typically less frequent, more contextually specialized, and acquired later — a pattern reflecting the inherent complexity of domain-specific vocabulary. Such cross-categorical integration suggests that lexical sophistication operates as a multidimensional construct with interconnected subcomponents.

The results of this study further corroborate the fundamental dichotomy between content and function words observed in prior research. Echoing Durrant and Brenchley's (2019) methodological distinction, the factor analysis demonstrates divergent clustering patterns: content word indices reflect lexical sophistication, whereas function word measures primarily correlate with syntactic complexity. This theoretical division finds empirical support in the genre comparison of this study, project proposals exhibited greater sophistication in content words but simpler function word usage compared to essays. This contrast aligns with their communicative purposes: specialized content vocabulary serves informational/persuasive goals in project proposals by establishing technical authority, while argumentative essays prioritize syntactic complexity through elaborated clausal structures to scaffold logical argumentation.

5.3. Lexical Sophistication in Project Proposal Quality Assessment

The diminished predictive power of lexical sophistication in the quality assessment of project proposals, as evidenced in this study, underscores the necessity of contextualizing writing evaluation frameworks within genre-specific demands. Three interrelated dimensions merit elaboration.

First, genre conventions fundamentally redefine how lexical features are evaluated. While lexical sophistication in argumentative essays may reflect discursive complexity and is often associated with critical thinking capabilities, project proposals place greater value on clarity and functional precision. Project proposals require writers to convey technical details (e.g., methodology, feasibility) in a clear, concise manner, using standardized terms and conventional phrases to maintain a professional tone. These genre-specific expectations are reflected in how proposals are assessed: evaluators tend to prioritize logical flow, methodological soundness, and structural clarity over stylistic richness.

Second, once learners reach a certain level of proficiency, lexical differences become harder to detect, as most writers are able to meet the basic expectations for vocabulary use. Participants' homogeneous language proficiency (CET-4/B1 to CET-6/B2) and English scientific training in their specialty indicate mastery of core disciplinary vocabulary, reducing inter-text variability in basic lexical features. At advanced proficiency levels, conventional lexical metrics become less effective in distinguishing quality, as writers prioritize clarity over linguistic embellishment. This aligns with Factor 1's weak correlation with scores, suggesting evaluators place greater emphasis on textual clarity than lexical sophistication in project proposal assessment.

Finally, function word contextualization (Factor 5) reflects genre-specific textual adaptation. Factor 5's weak yet significant correlation with project proposal

quality reflects the shared contribution of three related function word properties: concreteness, imageability, and contextual distinctiveness. For example, consider the phrase "the experiment was conducted *in controlled conditions*" versus "the experiment was conducted *under controlled conditions*." The latter choice (*under*) more clearly frames the environmental constraints as externally imposed and precisely defined, reducing ambiguity about the experimental setup. Spatial prepositions like *through* (as in "through iterative testing") help readers visualize the process, making the description more concrete and engaging. Therefore, strategic use of concrete, imageable, and contextually distinct function words directly enhances project proposal quality through balancing terminological precision with discourse adaptability, a critical yet understudied factor in assessment frameworks.

6. Pedagogical Implications

The findings of this study yield practical insights for teaching project proposal writing in ESP contexts. First, because lexical sophistication alone does not strongly predict proposal quality, instruction should target broader discourse features, such as coherence, logical structure, and audience awareness. Teaching students how to organize their ideas clearly and tailor content to the expectations of reviewers or stakeholders may be more effective than vocabulary exercises alone. Another important point concerns function words. The way these words are used in context can significantly affect how a proposal is received. Helping students notice subtle but meaningful choices, such as opting for more concrete or dynamic function words, can strengthen the link between vocabulary use and communicative effectiveness within the genre. Finally, incorporating genre-based teaching tools, like annotated sample proposals and structured peer review, can give students clearer insight into how word choices influence evaluations. By examining strong examples, learners can see how technical terms and carefully chosen function words work together to meet disciplinary expectations. This dual focus, on accuracy and rhetorical flexibility, can better prepare engineering students for the kinds of writing they will do in professional settings.

7. Conclusion

This study contributes to the existing research on lexical sophistication by examining it as a multidimensional construct in the context of an ESP genre, project proposals produced by EFL students. Comparative analyses revealed significant differences with large effect sizes in lexical sophistication indices between project proposals and argumentative essays, likely due to the unique communicative goals and vocabulary demands of the former. Stepwise regression analysis demonstrated that lexical sophistication indices held limited predictive power for the quality of project proposals, revealing that lexical features may play a marginal role in shaping overall quality assessments in specialized genres.

The study has a few limitations. First, focusing solely on EFL learners' project

proposals narrows the scope of lexical analysis as learner writing may not capture the full range of linguistic strategies typically found in professional proposals. As a result, the links observed between lexical features and proposal quality might reflect how learners compensate for limited proficiency, rather than patterns that are truly characteristic of the genre. Second, the sample size was relatively small and fairly homogeneous, which may have reduced variation in lexical usage. This could, in turn, have led to an underestimation of how well lexical sophistication predicts writing quality, especially for lower-proficiency writers. In addition, the analysis focused primarily on lexical features and did not account for discourse-level elements like rhetorical structure or cohesion between sentences, both of which are known to influence how texts are evaluated overall.

These limitations point to several directions for future work. Including a larger and more diverse set of writing samples could help determine whether the lexical patterns found here are specific to learners or reflect broader genre conventions. Adding a discourse-level perspective, for example, by analyzing rhetorical moves or cohesive devices, could also offer a more complete picture of what makes a strong proposal. Taken together, these steps would help bring automated analysis closer to how human raters actually judge writing, and in turn support more effective teaching and assessment of ESP genres.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Andrews, S. (1989). Frequency and Neighborhood Effects on Lexical Access: Activation or Search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802-814. https://doi.org/10.1037//0278-7393.15.5.802

Baese-Berk, M. M., Drake, S., Foster, K., Lee, D., Staggs, C., & Wright, J. M. (2021). Lexical Diversity, Lexical Sophistication, and Predictability for Speech in Multiple Listening Conditions. *Frontiers in Psychology, 12,* Article ID: 661415. https://doi.org/10.3389/fpsyg.2021.661415

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B. et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445-459. https://doi.org/10.3758/bf03193014

Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press. https://doi.org/10.1017/cbo9780511621024

BNC Consortium (2007). *British National Corpus*. http://www.natcorp.ox.ac.uk

Bulté, B., & Housen, A. (2012). Defining and Operationalising L2 Complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Language Learning & Language Teaching* (pp. 21-46). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.32.02bul

Cheng, A. (2025). Genre and ESP. In S. Starfield, & C. A. Hafner (Eds.), *The Handbook of English for Specific Purposes* (2nd ed., pp. 149-165). John Wiley & Sons. https://doi.org/10.1002/9781119985068.ch8

Coltheart, M. (1981). The MRC Psycholinguistic Database. *The Quarterly Journal of Experimental Psychology Section A, 33,* 497-505.

https://doi.org/10.1080/14640748108400805

Connor, U., & Mauranen, A. (1999). Linguistic Analysis of Grant Proposals: European Union Research Grants. *English for Specific Purposes, 18,* 47-62. https://doi.org/10.1016/s0889-4906(97)00026-4

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly, 34,* 213-238. https://doi.org/10.2307/3587951

Crossley, S. A., & Skalicky, S. (2019). Examining Lexical Development in Second Language Learners: An Approximate Replication of Salsbury, Crossley & Mcnamara (2011). *Language Teaching, 52,* 385-405. https://doi.org/10.1017/s0261444817000362

Crossley, S. A., Cai, Z., & McNamara, D. S. (2012). Syntagmatic, Paradigmatic, and Automatic N-Gram Approaches to Assessing Essay Quality. In P. M. McCarthy, & G. M. Youngblood (Eds.), *Proceedings of the 25th international Florida artificial intelligence Research Society (FLAIRS) Conference* (pp. 214-219). The AAAI Press.

Crossley, S., Salsbury, T., & McNamara, D. (2009). Measuring L2 Lexical Growth Using Hypernymic Relationships. *Language Learning, 59,* 307-334. https://doi.org/10.1111/j.1467-9922.2009.00508.x

Crossley, S., Salsbury, T., & McNamara, D. (2010). The Development of Polysemy and Frequency Use in English Second Language Speakers. *Language Learning, 60,* 573-605. https://doi.org/10.1111/j.1467-9922.2010.00568.x

Dascalu, M., McNamara, D., Crossley, S., & Trausan-Matu, S. (2016). Age of Exposure: A Model of Word Learning. *Proceedings of the AAAI Conference on Artificial Intelligence, 30,* 2928-2934. https://doi.org/10.1609/aaai.v30i1.10372

Durrant, P., & Brenchley, M. (2019). Development of Vocabulary Sophistication across Genres in English Children's Writing. *Reading and Writing, 32,* 1927-1953. https://doi.org/10.1007/s11145-018-9932-8

Durrant, P., Moxley, J., & McCallum, L. (2019). Vocabulary Sophistication in First-Year Composition Assignments. *International Journal of Corpus Linguistics, 24,* 33-66. https://doi.org/10.1075/ijcl.17052.dur

Eguchi, M., & Kyle, K. (2020). Continuing to Explore the Multidimensional Nature of Lexical Sophistication: The Case of Oral Proficiency Interviews. *The Modern Language Journal, 104,* 381-400. https://doi.org/10.1111/modl.12637

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database.* The MIT Press. https://doi.org/10.7551/mitpress/7287.001.0001

Field, A. (2018). *Discovering Statistics Using SPSS* (5th ed.). Sage.

Flowerdew, L. (2016). A Genre-Inspired and Lexico-Grammatical Approach for Helping Postgraduate Students Craft Research Grant Proposals. *English for Specific Purposes, 42,* 1-12. https://doi.org/10.1016/j.esp.2015.10.001

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,* 218-238. https://doi.org/10.1016/j.asw.2013.05.002

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). Pearson.

Higginbotham, G., & Reid, J. (2019). The Lexical Sophistication of Second Language Learners' Academic Essays. *Journal of English for Academic Purposes, 37,* 127-140. https://doi.org/10.1016/j.jeap.2018.12.002

Hyland, K. (2004). *Genre and Second Language Writing.* University of Michigan Press/ELT. https://doi.org/10.3998/mpub.23927

Jarvis, S. (2013). Defining and Measuring Lexical Diversity. In S. Jarvis, & M. Daller (Eds.), *Vocabulary Knowledge: Human Ratings and Automated Measures* (pp. 13-43). John Benjamins.

Juanggo, W. (2018). Investigating Lexical Diversity and Lexical Sophistication of Productive Vocabulary in the Written Discourse of Indonesian EFL Learners. *Indonesian Journal of Applied Linguistics, 8,* 38-48. https://doi.org/10.17509/ijal.v8i1.11462

Kanoksilapatham, B. (2015). Distinguishing Textual Features Characterizing Structural Variation in Research Articles across Three Engineering Sub-Discipline Corpora. *English for Specific Purposes, 37,* 74-86. https://doi.org/10.1016/j.esp.2014.06.008

Kim, M., Crossley, S. A., & Kyle, K. (2018). Lexical Sophistication as a Multidimensional Phenomenon: Relations to Second Language Lexical Proficiency, Development, and Writing Quality. *The Modern Language Journal, 102,* 120-141. https://doi.org/10.1111/modl.12447

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-Acquisition Ratings for 30,000 English Words. *Behavior Research Methods, 44,* 978-990. https://doi.org/10.3758/s13428-012-0210-4

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. (2015). Automatically Assessing Lexical Sophistication: Indices, Tools, Findings, and Application. *TESOL Quarterly, 49,* 757-786. https://doi.org/10.1002/tesq.194

Kyle, K., Crossley, S., & Berger, C. (2018). The Tool for the Automatic Analysis of Lexical Sophistication (TAALES): Version 2.0. *Behavior Research Methods, 50,* 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,* 307-322. https://doi.org/10.1093/applin/16.3.307

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,* Article 100894. https://doi.org/10.1016/j.jeap.2020.100894

Loewen, S., & Gonulal, T. (2015). Exploratory Factor Analysis and Principal Components Analysis. In L. Plonsky (Ed.), *Advancing Quantitative Methods in Second Language Research* (, pp. 182-212). Routledge. https://doi.org/10.4324/9781315870908-9

Lu, X., & Hu, R. (2021). Sense-Aware Lexical Sophistication Indices and Their Relationship to Second Language Writing Quality. *Behavior Research Methods, 54,* 1444-1460. https://doi.org/10.3758/s13428-021-01675-6

McDonald, S. A., & Shillcock, R. C. (2001b). Rethinking the Word Frequency Effect: The Neglected Role of Distributional Information in Lexical Processing. *Language and Speech, 44,* 295-322. https://doi.org/10.1177/00238309010440030101

McDonald, S., & Shillcock, R. (2001a). *Contextual Distinctiveness: A New Lexical Property Computed from Large Corpora*. (Informatics Research Report; No. EDI-INF-RR-0042). School of Informatics.

Nesi, H., & Gardner, S. (2012). *Genres across the Disciplines: Student Writing in Higher Education*. Cambridge University Press. https://doi.org/10.1017/9781009030199

Plonsky, L., & Gonulal, T. (2015). Methodological Synthesis in Quantitative L2 Research: A Review of Reviews and a Case Study of Exploratory Factor Analysis. *Language Learning, 65,* 9-36. https://doi.org/10.1111/lang.12111

Read, J. (2000). *Assessing Vocabulary*. Cambridge University Press. https://doi.org/10.1017/cbo9780511732942

Salsbury, T., Crossley, S. A., & McNamara, D. S. (2011). Psycholinguistic Word Information in Second Language Oral Discourse. *Second Language Research, 27,* 343-360. https://doi.org/10.1177/0267658310395851

Simpson-Vlach, R., & Ellis, N. C. (2010). An Academic Formulas List: New Methods in Phraseology Research. *Applied Linguistics, 31,* 487-512. https://doi.org/10.1093/applin/amp058

Sun, K., & Lu, X. (2021). Assessing Lexical Psychological Properties in Second Language Production: A Dynamic Semantic Similarity Approach. *Frontiers in Psychology, 12,* Article ID: 672243. https://doi.org/10.3389/fpsyg.2021.672243

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings.* Cambridge University Press.

Tracy-Ventura, N. (2017). Combining Corpora and Experimental Data to Investigate Language Learning during Residence Abroad: A Study of Lexical Sophistication. *System, 71,* 35-45. https://doi.org/10.1016/j.system.2017.09.022

Vögelin, C., Jansen, T., Keller, S. D., Machts, N., & Möller, J. (2019). The Influence of Lexical Features on Teacher Judgements of ESL Argumentative Essays. *Assessing Writing, 39,* 50-63. https://doi.org/10.1016/j.asw.2018.12.003

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory Factor Analysis: A Five-Step Guide for Novices. *Australasian Journal of Paramedicine, 8,* 1-13. https://doi.org/10.33151/ajp.8.3.93

Zhang, X., Lu, X., & Li, W. (2021). Beyond Differences: Assessing Effects of Shared Linguistic Features on L2 Writing Quality of Two Genres. *Applied Linguistics, 43,* 168-195. https://doi.org/10.1093/applin/amab007

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

Zhou, X., Gao, Y., & Lu, X. (2023). Lexical Complexity Changes in 100 Years' Academic Writing: Evidence from Nature Biology Letters. *Journal of English for Academic Purposes, 64,* Article 101262. https://doi.org/10.1016/j.jeap.2023.101262