Journal of new advances in English Language Teaching and Applied Linguistics (JELTAL)
Summer-Autumn 7 (2) 2025

Research paper

Evaluating the Predictive Power of Fine-Grained Syntactic Complexity Measures for IELTS Writing Scores

Ehsan Abbaspour[1]*, Priya Mathew[1]

[1]*Centre for Academic Writing, Middle East College, Muscat, Omann*

Abbaspour, E, & Mathew, P. (2025).Evaluating the predictive power of fine-grained syntactic complexity measures for IELTS writing scores. *Journal of new advances in English Language Teaching and Applied Linguistics*, *7*(2), 89-112.
Doi:10.22034/Jeltal.2025.7.2.5

*corresponding author: *ehsan@mec.edu.om*

Abstract

The present study investigates fine-grained syntactic complexity measures and their potential role in predicting the quality of L2 writing in IELTS Writing Task 2. A corpus of 105 IELTS essays, covering a wide range of band scores for Grammatical Range and Accuracy (GRA), was systematically analyzed using the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC). In total, thirty-one clausal and eighteen phrasal syntactic indices were examined in detail. Correlation analyses revealed that several indices, such as passive constructions and nominal phrase dependents, serve as strong predictors of GRA scores. In particular, indices linked to noun phrase elaboration and passive structures demonstrated significant positive associations with higher GRA bands, suggesting that these features contribute to more advanced and sophisticated perceptions of grammatical complexity. Overall, these findings reinforce previous research and highlight that incorporating fine-grained syntactic complexity into writing assessments can improve scoring accuracy, strengthen objectivity, and enhance the reliability of automated evaluations.

Keywords: Syntactic complexity; Phrasal complexity; Clausal complexity; L2 writing quality

Received: *2025-05-22* **Revised** *2025-09-10* **Accepted** *2025-09-18*

Introduction

Within the IELTS Writing Task 2, the criterion of Grammatical Range and Accuracy (GRA) serves as a key indicator of a candidate's linguistic competence, reflecting their ability to construct

varied and sophisticated sentence structures while maintaining grammatical accuracy (IELTS, n.d.). This aspect of writing performance is not only integral to achieving higher scores but also serves as a reliable predictor of overall language proficiency and academic writing ability (Banerjee et al., 2007; Kang et al., 2023; Müller & Han, 2022).

The investigation of valid and reliable indicators of syntactic complexity that can effectively predict second language (L2) writing quality has been a focal point of research in recent decades (e.g., Biber et al., 2016; Crossley, 2020; Crossley & McNamara, 2014; Kim &Crossley, 2018; Kyle & Crossley, 2018; Seo & Oh, 2024). Historically, holistic T-unit-based and clausal measures of syntactic complexity, which do not distinguish between subtypes of clausal or phrasal structures (e.g., T-units per sentence, clauses per T-unit, and dependent clauses per clause), have been central to studies examining L2 writing development and quality (Lu, 2011, 2017; Ortega, 2003; Wolfe–Quintero et al., 1998). Hunt (1965, p.21) defines T-units as "the shortest grammatically allowable sentences into which the theme could be segmented".

However, a more recent strand of research, largely influenced by usage-based approaches to second language acquisition (Kyle et al., 2021), has shifted focus towards more fine-grained clausal and phrasal indices. These measures account for the structural types of clauses (e.g., clausal prepositional complements per clause and clausal subjects per clause) and phrases (e.g., adjectives per noun phrase as direct object and prepositions per noun phrase as nominal subject) (e.g., Biber et al., 2016; Kim & Crossley, 2018).

Despite these advancements, discrepancies persist in research comparing the relative predictive power of traditional versus fine-grained syntactic complexity indices for L2 writing quality (e.g., Kim & Crossley, 2018). Moreover, this predictive power has yet to be examined specifically in the context of IELTS writing, a significant gap in the literature given the widespread use of this

high-stakes test in evaluating English language proficiency for academic and professional purposes.

This research aims to address the gap by conducting a comprehensive analysis of fine-grained syntactic complexity measures in relation to L2 writing quality, with a particular focus on IELTS writing. Such an investigation promises to yield valuable insights into the most effective indicators of syntactic complexity in L2 writing assessment.

Literature Review

Traditional (Large-grained) Measures of Syntactic Complexity

A wide range of quantitative measures have been used to assess syntactic complexity in L2 writing. These differ in terms of their focus and granularity. However, three general categories can be identified: length-based measures, coordination measures, and subordination measures (Wolfe-Quintero et al., 1998).

Length-based measures simply calculate the average length of particular syntactic units such as words, clauses, or T-units. The most commonly used indices are mean length of clause (MLC), mean length of T-unit (MLTU), and mean length of sentence (MLS). These provide general indicators of syntactic complexity but reveal little about the specific structures writers are using.

Coordination measures quantify the amount of coordination through indices like coordinate phrases per clause (CP/C) and coordinate phrases per T-unit (CP/T) (Bardovi-Harlig, 1992; Lu, 2011). These capture learners' ability to conjoin syntactic units using coordinating conjunctions like "and" and "but".

Subordination indices assess clausal complexity through clauses per T-unit (C/TU), dependent clauses per clause (DC/C) and dependent clauses per T-unit (DC/T) (Lu, 2011; Ortega, 2003).

These ratios reflect how frequently writers elaborate T-units through dependent adverbial and noun clauses.

Such indices can be analyzed using computational tools that automatically annotate texts for syntactic structures. For example, Lu's (2010) L2 Syntactic Complexity Analyzer (L2SCA) annotates 14 traditional measures like MLT and also tags texts for the frequency of structures including complex nominals, coordinate phrases, and particular clausal types. These tagging programs allow efficient calculation of fine-grained indices.

While easily calculated, length-based and subordination measures have been critiqued for their lack of specificity regarding the particular structures written (Biber et al., 2011; Norris & Ortega, 2009).

Fine-grained Phrasal and Clausal Measures of Syntactic Complexity

While traditional length-based indices provide useful overall indicators of syntactic complexity, researchers have increasingly recognized the need for more fine-grained measures capable of capturing specific structures (Bulté&Housen, 2012; Lu, 2011). Fine-grained measures aim to quantify particular syntactic features and constructions to yield more detailed and meaningful profiles of learner language. In L2 writing specifically, fine-grained measures provide valuable insights into developing syntactic skills. For instance, Kyle and Crossley (2015) discovered indices relating to phrase-level modification were stronger predictors of essay scores than traditional MLTU and clause ratios. This reinforces the idea that academic writing development entails expanding noun phrase complexity.

More precise profiling requires manual annotation by trained coders. Biber et al. (2011) manually annotated conversational and academic texts for over 70 fine-grained features such as

dependent clauses, phrasal modifiers, and passive constructions. This intensive process yields comprehensive syntactic profiles but is extremely time-consuming, limiting corpus sizes.Nonetheless, limitations exist. Many structures occur infrequently, making robust corpus analysis difficult (Ravid & Berman, 2010) and manual annotation of large corpora is extremely laborious.

To mitigate some of these limitations, some solutions have been proposed and tools developed to analyze the texts more rapidly and accurately. To balance precision and efficiency, some researchers have proposed selective manual annotation of structures resistant to accurate computational tagging, supplemented by automated analysis of other features (Biber & Gray, 2016). For instance, passive constructions may require manual coding while other indices can be automatically extracted. This combined approach leverages the strengths of both human coders and computational tools.

One of the tools that allows the automatic extraction of syntactic structures is the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) developed by Kyle (2016). Unlike Lu's (2011) L2SCA, which incorporates traditional measures, TAASSC (Kyle, 2016; Kyle & Crossley, 2018) uses 31 different indices to calculate fine-grained clausal complexity. To mitigate bias towards structures that contain more words, TAASSC assesses clause length by measuring the number of direct dependents per clause rather than the word count. Additionally, it counts each type of structure separately rather than grouping them. It also considers a clause to be composed of a primary verb and any accompanying structures, which may or may not contain a finite verb.

Phrasal indices for 10 different types of phrasal dependents and seven different types of noun phrases are also included in TAASSC. Phrasal indices fall into three groups. The average number

of dependents for each particular phrase type as well as for all phrase types is calculated in the first category. Regardless of the sort of noun phrase they follow, the second group counts the frequency of particular dependent types. The average occurrence of certain dependent types within particular types of noun phrases is determined by the third category of phrasal indices. In English, noun phrases may include pronouns, which typically do not have direct dependents. To accommodate this, TAASSC includes two versions of each index: one that incorporates pronoun noun phrases into its computations and another that does not. The average number of phrasal dependents per phrase type, such as the average number of dependents per nominal subject, is indicated by the basic phrasal indices in TAASSC.

Several recent studies have employed TAASSC to analyze fine-grained indices of syntactic complexity in writing. Li et al. (2023) examined the predictive power of fine-grained syntactic complexity (SC) indices for letter writing proficiency and their relationship to pragmatic appropriateness in Chinese adolescent learners of English as a foreign language. To this end, 300 request letters and 300 self-recommendation letters written by these learners were analyzed focusing on 10 fine-grained indices related to dependent clause (DC) and complex noun phrase (CNP) usage using TAASSC. The study also investigated the correlation between significant SC predictors for proficiency ratings and the degrees of politeness and formality in the letters. The results indicated that the clausal and noun phrase features associated with the fine-grained DC- and CNP-related indices were predictive of the rated proficiency levels of the letters and reflected the varied degrees of politeness and formality employed by learners of different English writing proficiency.

In another study, Zhang and Ouyang (2023) investigated the predictive power of syntactic complexity, as measured by TAASSC in assessing the writing quality in independent and

integrated writing tasks. The researchers collected writing samples from intermediate EFL learners and used regression models to examine the relationship between syntactic complexity indices and the ratings of different writing tasks. The main findings of the study indicated that syntactic complexity, particularly fine-grained indices related to phrasal and clausal complexity, had significant predictive power over the writing quality of both independent and integrated writing tasks.

Li and Yang (2023) investigated the relationship between syntactic complexity and writing quality in research articles. The researchers also used TAASSC to analyze various syntactic indices in different sections of the research articles, including the abstract, introduction, and method. The main findings of the study revealed that high-quality articles generally had incorporated more complex structures compared to their low-quality counterparts. The study also found that a combination of absolute (i.e., traditional) and relative (i.e., fine-grained) measures provided the best predictive results for assessing the writing quality of articles.

The cumulative findings of these studies demonstrate the complicated nature of syntactic complexity in L2 writing. The results support a more sophisticated method of assessing and teaching L2 writers, one that goes beyond traditional indices and takes meaning-based and functional features of syntactic complexity into account.

Also, in the studies reviewed above, the relationship between syntactic complexity and overall writing scores was examined. The overall writing score encompasses various factors, including the quality of content, text organization, and the accuracy and complexity of grammatical and lexical features. Thus, there is a need for comparing syntactic complexity indices with human assessments of texts, focusing on grammatical features alone.

Currently, IELTS Writing essays are evaluated by trained human examiners using the criteria established by the IELTS. The examiners undergo an extensive and rigorous training process, which is supplemented by periodic retraining sessions to maintain scoring standards and consistency. This process is both time-consuming and costly. Therefore, developing methods to objectively quantify and analyze such structures could significantly improve the efficiency of scoring essays in this high-stakes test. Accordingly, the following research questions were formulated:

1. Which fine-grained clausal syntactic complexity measures correlate with human examiners' scores for Grammatical Range and Accuracy (GRA) in IELTS Writing Task 2?
2. Which fine-grained phrasal syntactic complexity measures correlate with human examiners' scores for Grammatical Range and Accuracy (GRA) in IELTS Writing Task 2?

Methodology

Corpus

The data in this study consists of 105 IELTS Writing Task 2 essays written on various prompts and scored by IELTS-trained official examiners. These essays were sourced from official IELTS samples published by Cambridge University, the British Council, and IDP Australia, the three co-owners of the test. The researchers corrected spelling errors in the essays to prevent data contamination, as spelling is evaluated under Lexical Resources in IELTS rather than Grammatical Range and Accuracy. As depicted in Table 1, the sample essays span a wide range of band scores for Grammatical Range and Accuracy.

Table 1
The Scored Essays Used as the Corpus of the Study

Band Score	Number of Sample Essays	Word Count
9	22	6048
8	9	2596
7	27	6864
6	27	7849
5	16	3677
4	4	858
Total	105	27892

Table 2 presents the band descriptors for Grammatical Range and Accuracy across different bands in IELTS Writing Task 2.

Table 2
IELTS Writing Task 2 Band Score Descriptors for Grammatical Range and Accuracy

GRA Score	Descriptors
9	The candidate uses a wide range of structures with full flexibility and accuracy. Rare minor errors occur only as "slips."
8	The candidate uses a wide range of structures effectively. Most sentences are error-free, with only occasional inaccuracies or inappropriacies.
7	The candidate uses a variety of complex structures. Frequent error-free sentences are evident, though some errors persist in more complex structures.
6	The candidate uses a mix of simple and complex sentence forms. While errors occur, they rarely impede communication.
5	The candidate uses only a limited range of structures. Frequent errors in grammar and punctuation may cause some difficulty for the reader.
4	The candidate has limited control of simple structures and often makes errors, which may hinder communication.

Measures of Syntactic Complexity

The study incorporated fine-grained indices of syntactic complexity, 31 of which were clausal indices (Table 3) and 18, phrasal (Table 4). The indices were drawn from Kyle and Crossley (2018).

Table 3

Fine-grained Clausal Syntactic Complexity Indices

Abbreviation	Description
cl_av_deps	dependents per clause
acomp_per_cl	adjective complements per clause
advcl_per_cl	adverbial clauses per clause
agent_per_cl	passive agents per clause
cc_per_cl	clausal coordinating conjunctions per clause
ccomp_per_cl	clausal complements per clause
conj_per_clause	conjunctions per clause
csubj_per_cl	clausal subjects per clause
csubjpass_per_cl	passive clausal subjects per clause
dep_per_cl	undefined dependents per clause
discourse_per_cl	discourse markers per clause
dobj_per_cl	direct objects per clause
expl_per_cl	existential "there" per clause
iobj_per_cl	indirect objects per clause
mark_per_cl	subordinating conjunctions per clause
ncomp_per_cl	nominal complements per clause
neg_per_cl	clausal negations per clause
nsubj_per_cl	nominal subjects per clause
nsubjpass_per_cl	passive nominal subjects per clause
parataxis_per_cl	instances of parataxis per clause
pcomp_per_cl	clausal prepositional complements per clause
prep_per_cl	prepositions per clause
prepc_per_cl	prepositional clausal modifiers per clause
prt_per_cl	phrasal verb particles per clause
tmod_per_cl	bare noun phrase temporal modifiers per clause
xcomp_per_cl	open clausal complements per clause
xsubj_per_cl	controlling subjects per clause
advmod_per_cl	adverbial modifiers per clause
aux_per_cl	auxiliary verbs per clause
auxpass_per_cl	passive auxiliary verbs per clause
modal_per_clause	modal auxiliaries per clause

Table 4

Fine-grained Phrasal Syntactic Complexity Indices

Abbreviation	Description
av_nominal_deps	dependents per nominal
av_nsubj_deps	dependents per nominal subject
av_nsubj_pass_deps	dependents per passive nominal subject
av_agents_deps	dependents per agent
av_dobj_deps	dependents per direct object
av_pobj_deps	dependents per object of the preposition
av_iobj_deps	dependents per indirect object

av_ncomp_deps	dependents per nominal complement
det_all_nominal_deps_struct	determiners per nominal
amod_all_nominal_deps_struct	adjectival modifiers per nominal
prep_all_nominal_deps_struct	prepositions per nominal
poss_all_nominal_deps_struct	possessives per nominal
vmod_all_nominal_deps_struct	verbal modifiers per nominal
nn_all_nominal_deps_struct	nouns as a nominal dependents per nominal
rcmod_all_nominal_deps_struct	relative clause modifiers per nominal
advmod_all_nominal_deps_struct	(non-clausal) adverbial modifiers per nominal
conj_and_all_nominal_deps_struct	conjunction "and" as a nominal dependent per nominal
conj_or_all_nominal_deps_struct	conjunction "or" as a nominal dependent per nominal

Data Analysis

In this study, fine-grained indices (phrasal and clausal) were assessed using TAASSC (Kyle, 2016). Given the non-normal distribution of the data for GRA band scores (Table 5), Spearman's Rank Correlation was employed using SPSS to examine the relationship between human holistic scoring of IELTS Task 2 essays and both traditional and fine-grained indices of syntactic complexity.

Table 5
Test of Normality for GRA Band Scores

	Shapiro-Wilk		
	Statistic	df	Sig.
GRA Band Score	.909	105	.000

Note. The Sig. value below 0.05 indicates deviation from a normal distribution.

Results

Research Question 1

The first research question examined the extent of the correlation between holistic scoring of IELTS essays for GRA and fine-grained indices of syntactic complexity at the clausal level. To address this TAASSC was used to calculate the clausal indices of syntactic complexity for the

essays. Subsequently, a Spearman's Rank Correlation was conducted using SPSS to analyze the relationship between these indices and the GRA bands awarded by the examiners. The results are presented in Table 6.

Table 6
Correlation between Holistic Human Scoring and Clausal Indices of Syntactic Complexity

Variable	Correlation Coefficient (*r*)	Sig. (2-tailed) *p*	*N*
Dependents per clause	.016	.874	105
Adjective complements per clause	.223*	.022	105
Adverbial clauses per clause	.017	.863	105
Passive agents per clause	.375**	.000	105
Clausal coordinating conjunctions per clause	-.158	.108	105
Clausal complements per clause	-.036	.718	105
Conjunctions per clause	-.030	.762	105
Clausal subjects per clause	.284**	.003	105
Passive clausal subjects per clause	-.057	.564	105
Undefined dependents per clause	-.071	.473	105
Discourse markers per clause	-.110	.263	105
Direct objects per clause	-.257**	.008	105
Existential "there" per clause	.121	.221	105
Indirect objects per clause	.027	.786	105
Subordinating conjunctions per clause	.159	.105	105
Nominal complements per clause	-.011	.911	105
Clausal negations per clause	-.024	.806	105
Nominal subjects per clause	-.270**	.005	105
Passive nominal subjects per clause	.435**	.000	105
Instances of parataxis per clause	.028	.777	105
Clausal prepositional complements per clause	-.057	.564	105
Prepositions per clause	.081	.410	105
Prepositional clausal modifiers per clause	-.096	.332	105
Phrasal verb particles per clause	-.035	.725	105
Bare noun phrase temporal modifiers per clause	-.021	.835	105
Open clausal complements per clause	.010	.918	105
Controlling subjects per clause	-	-	105
Adverbial modifiers per clause	.160	.104	105
Auxiliary verbs per clause	-.046	.639	105
Passive auxiliary verbs per clause	.432**	.000	105
Modal auxiliaries per clause	.147	.134	105

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

A range of indices demonstrated statistically significant correlations with GRA scores, with both positive and negative associations observed. Variables positively correlated with GRA scoring included "Adjective complements per clause" (r = .223, p< .05), "Passive agents per clause" (r = .375, p< .01), "Clausal subjects per clause" (r = .284, p< .01), "Passive nominal subjects per clause" (r = .435, p< .01), and "Passive auxiliary verbs per clause" (r = .432, p< .01).

Conversely, several indices demonstrated a significant negative correlation with GRA scores, indicating an inverse relationship. These included "Direct objects per clause" (r = −.257, p< .01) and "Nominal subjects per clause" (r = −.270, p< .01). Additionally, a range of clausal indices did not demonstrate statistically significant correlations with GRA scores. These indices included "Dependents per clause," "Adverbial clauses per clause," "Clausal coordinating conjunctions per clause," "Clausal complements per clause," "Conjunctions per clause," "Passive clausal subjects per clause," "Undefined dependents per clause," "Discourse markers per clause," "Existential 'there' per clause," "Indirect objects per clause," "Subordinating conjunctions per clause," "Nominal complements per clause," "Clausal negations per clause," "Instances of parataxis per clause," "Clausal prepositional complements per clause," "Prepositions per clause," "Prepositional clausal modifiers per clause," "Phrasal verb particles per clause," "Bare noun phrase temporal modifiers per clause," "Open clausal complements per clause," "Adverbial modifiers per clause," "Auxiliary verbs per clause," and "Modal auxiliaries per clause. “Finally, Controlling subjects per clause" showed no recorded instances in the corpus, thus no correlation could be assessed for this index.

Research Question 2

The second research question investigated the potential relationship between holistic scoring of IELTS essays for GRA and fine-grained indices of syntactic complexity at the phrasal level. For this purpose, a Spearman's Rank Correlation was performed using the output data from TAASSC for the IELTS essays and the GRA band scores awarded by human examiners. The results are illustrated in Table 7.

Table 7
Correlation between Holistic Human Scoring and Phrasal Indices of Syntactic Complexity

Variable	Correlation Coefficient (*r*)	Sig. (2-tailed) *p*	*N*
Dependents per nominal	.378**	.000	105
Dependents per nominal subject	.342**	.000	105
Dependents per passive nominal subject	.267**	.006	105
Dependents per agent	.373**	.000	105
Dependents per direct object	.281**	.004	105
Dependents per object of the preposition	.245*	.012	105
Dependents per indirect object	-.134	.172	105
Dependents per nominal complement	.275**	.005	105
Determiners per nominal	.076	.439	105
Adjectival modifiers per nominal	.423**	.000	105
Prepositions per nominal	.406**	.000	105
Possessives per nominal	.190	.052	105
Verbal modifiers per nominal	.040	.683	105
Nouns as a nominal dependents per nominal	.208*	.033	105
Relative clause modifiers per nominal	-.102	.299	105
(Non-clausal) adverbial modifiers per nominal	.178	.070	105
Conjunction "and" as a nominal dependent	.328**	.001	105
Conjunction "or" as a nominal dependent	.241*	.013	105

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

The analysis revealed that nine indices exhibited strong positive correlations with GRA scores at the 0.01 level. The indices with the highest correlations included “Adjectival modifiers per nominal” (r = .423, p = .000) and “Prepositions per nominal” (r = .406, p = .000), each with

correlation coefficients above .4, indicating a strong positive association with GRA scores. Other indices with strong correlations were “Dependents per nominal” (r = .378, p = .000), “Dependents per agent” (r = .373, p = .000), “Dependents per nominal subject” (r = .342, p = .000), “Conjunction *and* as a nominal dependent” (r = .328, p = .001), “Dependents per nominal complement” (r = .275, p = .005), “Dependents per direct object” (r = .281, p = .004), and “Dependents per passive nominal subject” (r = .267, p = .006).

Three additional indices showed moderate positive correlations with GRA scores at the 0.05 level. These included “Dependents per object of the preposition” (r = .245, p = .012), “Nouns as a nominal dependents per nominal” (r = .208, p = .033), and “Conjunction *or* as a nominal dependent” (r = .241, p = .013).

In contrast, six indices were not significantly correlated with GRA scores. These included “Dependents per indirect object” (r = -.134, p = .172), “Determiners per nominal” (r = .076, p = .439), “Possessives per nominal” (r = .190, p = .052), “Verbal modifiers per nominal” (r = .040, p = .683), “Relative clause modifiers per nominal” (r = -.102, p = .299), and “(Non-clausal) adverbial modifiers per nominal” (r = .178, p = .070).

Discussion

Research Question 1

Regarding the first research question, an evaluation of fine-grained clausal syntactic complexity measures further elucidated how well these indices may reflect overall GRA scores. Positive correlations between GRA scores and syntactic indices, such as "Passive agents per clause," "Passive nominal subjects per clause," and "Adjective complements per clause," underscore the association between passivization, lexical modification, and higher-rated writing

(See Table 8). This is in line with studies by Kyle and Crossley (2018), who found that passivization frequently serves as a marker of advanced academic writing and contributes to a formal, impersonal tone. Furthermore, Zhang and Lu (2022) retain this view and state that the majority of adjective complements and complexes of nouns help achieve high GRA scores because they add to the syntactic and lexical complexity in argumentative tasks.

Table 8
Examples of Clausal Indices

Clausal Index	Examples
Adjective complements per clause	Many students find the academic workload **challenging** [acomp].
Passive agents per clause	The new policy was implemented **by the government** [agent].
Clausal subjects per clause	**That people are living longer** [csubj] is a result of advancements in healthcare.
Passive nominal subjects per clause	**The proposal** [nsubjpass] was accepted by the committee.
Passive auxiliary verbs per clause	The new regulations **were** [auxpass] implemented last year.

"Adjective complements per clause" was positively correlated with GRA, reinforcing the idea that adjective use enhances perceived grammatical range (Li et al., 2023). This employment of descriptive structures supports the results of Biber et al. (2016), as well as Zhang and Lu (2022), who stated that the ability to include adjectival modifiers greatly enhances the writing proficiency in a second language in all the genres. Additionally, "Passive agents per clause" and "Passive nominal subjects per clause" positively correlate with GRA scores, as these structures signal grammatical sophistication. According to Li et al. (2023) and Zhang and Lu (2022), for instance, passives create a distanced tone typical of many academic evaluations. In this regard, Seo and Oh (2024) also note that advanced L2 writers use passive voice constructions as one of the proficiency markers.

On the other hand, negative correlations were observed with indices such as "Direct objects per clause" and "Nominal subjects per clause", indicating that simpler syntactic structures might detract from perceived grammatical range. This trend is consistent with Kyle and Crossley (2018), who found that L2 writers with higher proficiency scores in TOEFL essays favored more embedded and complex noun phrases over direct object constructions. Such avoidance of simpler syntactic forms may reflect a stylistic preference for structures that align with academic writing norms, as noted by Seo and Oh (2024), who suggested that reduced reliance on basic clausal structures like direct objects can enhance perceived grammatical sophistication.

Furthermore, several clausal indices showed no statistically significant correlation with GRA scores, including "Adverbial clauses per clause" and "Subordinating conjunctions per clause." This lack of correlation indicates a limited role for clausal subordination in predicting GRA, supporting the argument put forth by Biber et al. (2011) that advanced L2 writers rely more on nominal complexity than clausal subordination in academic contexts. Zhang and Lu (2022) corroborate this perspective, noting that in argumentative writing, features such as adverbial and subordinate clauses may be less valued by examiners, who prioritize conciseness and phrasal sophistication over extensive clausal subordination. Similarly, Seo and Oh (2024) noted that advanced L2 writers are less reliant on clause-heavy structures, potentially due to the clarity and precision that nominal complexity affords in academic genres.

The findings thus align with the developmental framework proposed by Seo and Oh (2024), which suggests a progression from simpler to more sophisticated phrasal structures as markers of syntactic maturity among higher proficiency L2 writers. Zhang and Lu (2022) and Kyle and Crossley (2018) support this framework, observing that phrasal and nominal sophistication align more closely with writing quality ratings, while basic clausal structures may detract from perceived

complexity. This study reinforces the view that high-proficiency L2 writers benefit from adopting syntactic structures that enhance formality and lexical depth rather than relying on basic clausal constructions.

Overall, the results of this study suggest that perceived L2 writing proficiency is best enhanced through passive structures and nominal complexity, with a reduced focus on simpler clausal and subordinate elements. This shift toward phrasal over clausal complexity, as documented across sources, highlights the importance of targeted syntactic instruction and assessment practices that foster academic writing skills reflective of higher L2 proficiency.

Research Question 2

The investigation into fine-grained phrasal syntactic complexity measures revealed several significant correlations with GRA scores (See Table 9). Phrasal complexity indices, particularly those associated with noun phrase elaboration, consistently emerge as predictors of perceived writing quality, aligning with the findings of Biber et al. (2011) and Kyle and Crossley (2018). Specifically, indices such as Adjectival modifiers per nominal and Prepositions per nominal strongly correlate with GRA scores, suggesting that noun phrase elaboration significantly enhances syntactic sophistication in L2 academic writing (Kyle & Crossley, 2018; Taguchi et al, 2013). These results support Kyle and Crossley's (2018) conclusion that such phrasal elements may be stronger indicators of writing quality than clausal measures, particularly within formal and academic contexts.

Table 9
Examples of Phrasal Indices

Phrasal Index	Example Sentence
Dependents per nominal	**The recent government decision** [nominal] has sparked a significant debate.
Dependents per nominal subject	**Governments** [nsubj] should invest in **renewable** [dep] **energy** [dep] **sources** [nsubj].
Dependents per passive nominal subject	**The new educational policies** [nsubjpass] were introduced last year.
Dependents per agent	**The new traffic laws[dep]** were enforced **by city officials [agent].**
Dependents per direct object	Many people believe that **higher** [dep] **education** [dobj] should be a priority for the government.
Dependents per object of the preposition	Governments should invest more in **public [dep] transportation** [pobj].
Dependents per nominal complement	**The success of the program** [ncomp] was the result of careful planning.
Adjectival modifiers per nominal	**Affordable** [amod] **housing** is crucial for urban development.
Prepositions per nominal	**The impact of technology** [prep] **on education** [prep] is significant.
Nouns as a nominal dependents per nominal	The **university** [nn] **exam** [nn] system
Conjunction "and" as a nominal dependent	The **work and life** [conj_and] balance issues affect many professionals in urban areas.
Conjunction "or" as a nominal dependent	The **pass or fail** [conj_and] grading system has been debated in educational circles.

In addition to noun phrase modifiers, indices like Dependents per nominal and Dependents per direct object demonstrate predictive value, indicating that noun phrase density is a crucial marker of writing proficiency. The significant association between GRA scores and these indices underscores the importance of syntactically complex noun phrases, which Seo and Oh (2024) argue represent a developmental shift from simpler to more sophisticated structures in advanced L2 writing. Similarly, Crossley and McNamara (2014) found that noun phrases with multiple dependents are indicative of syntactic control and lexical depth, both essential for improved GRA scores.

Interestingly, indices like Conjunction “and” as a nominal dependent and Dependents per agent also correlate with higher GRA scores, suggesting that dense nominal constructions contribute to human evaluators' perceptions of grammatical range. Li et al. (2023) noted a similar preference for such complexity in adolescent EFL learners’ writing, where phrasal modifiers enhance the perceived specificity and formality of the text. The presence of such dense noun phrases aligns with the findings of Biber et al. (2016) and Zhang and Lu (2022), who argue that complex noun phrases (CNPs) are particularly influential in formal writing due to their role in clarifying meaning and adding academic tone.

Conversely, simpler indices like Determiners per nominal and Possessives per nominal show limited correlation with GRA scores, suggesting these features may have minimal impact on perceived grammatical range in IELTS essays. This finding aligns with Zhang and Lu’s (2022) observation that simpler structures may not carry sufficient evaluative weight in academic writing contexts, where human examiners often value phrasal complexity over basic nominal elements. Likewise, the lack of significant correlation for clausal complexity measures, such as Relative clause modifiers per nominal and Verbal modifiers per nominal, supports the notion by Biber et al. (2011) that L2 academic writing depends less on clausal structures and more on noun phrase elaboration. Seo and Oh (2024) similarly emphasize that academic genres tend to prioritize phrasal structures, especially within noun phrases, to maintain the precision and clarity expected in formal writing.

Furthermore, the correlation between Adjectival modifiers per nominal and GRA scores also supports findings from studies by Lan et al. (2019) and Kim (2020), who identified adjectival modifiers as strong indicators of syntactic maturity and grammatical range. Baik (1994) demonstrated a similar trend among Korean L1 writers, observing that as proficiency increases,

writers tend to shift toward noun phrase elaboration rather than clausal subordination. This syntactic density observed in high-proficiency writing, represented by indices such as Dependents per nominal subject and Prepositions per nominal, is valued by human evaluators, who associate these features with more sophisticated language use (Kim, 2020; Crossley & McNamara, 2014).

In summary, the study's findings reinforce the critical role of phrasal complexity indices, particularly noun-modifying elements, as indicators of grammatical range and accuracy. Targeted instructional practices that prioritize the development of these specific syntactic features—such as adjectival and prepositional modifiers—could help L2 learners achieve the syntactic sophistication expected in academic writing assessments.

Conclusion

This study aimed to investigate the predictive power of fine-grained syntactic complexity measures for evaluating L2 writing quality in the context of IELTS Writing Task 2. By analyzing a corpus of 105 essays with the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC), we examined thirty-one clausal and eighteen phrasal syntactic indices. The findings demonstrated that specific clausal and phrasal indices, such as passive agents per clause and dependents per nominal, are significant predictors of Grammatical Range and Accuracy (GRA) scores. These results suggest that the integration of syntactic complexity measures can enhance the objectivity and efficiency of scoring high-stakes language proficiency tests.

However, there are some limitations in the study. The relatively small corpus size has limited the generalizability of the findings. Moreover, the study solely accounted for syntactic complexity, whereas the IELTS scoring for GRA also includes accuracy, which was not considered in this study. Another limitation of the study is that the essays were written on different topics, which

may have influenced the writers' use of various grammatical structures. Future studies should be based on a far larger corpus with an equal number of samples from every band score to identify more precisely the indices corresponding to each band. Also, potentials of other automated analysis tools could be investigated for other IELTS writing criteria such as Cohesion and Coherence and Lexical Resources.

The findings have implications for IELTS preparatory courses, where structures that contribute more to higher GRA scores could be highlighted during grammar-focused sessions. The study also has implications for the development of an automated essay scoring system, which could make the assessment process of essays more objective, faster, and more cost-effective. This would align IELTS with other high-stake tests, such as TOEFL and PTE, which already use automated scoring for the writing section. Adopting automated systems may relieve some of the burden on human raters (even if they do not completely take over the scoring task) by letting them concentrate more on evaluating the content, which would in turn improve the overall efficiency and reliability of language proficiency evaluations.

REFERENCES

Baik, M. J. (1994). Syntactic features of Englishization in Korean. *World Englishes, 13*(2), 155-166.

Banerjee, J., Franceschina, F., & Smith, A. M. (2007). Documenting features of written language production typical at different IELTS band score levels. *IELTS Research Reports, 7*, 1-55.

Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly, 26*, 390-395. https://doi.org/10.2307/3587016

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511920776

Biber, D., Gray, B., &Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35. https://doi.org/10.5054/tq.2011.244483

Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics, 37*, 639-668. https://doi.org/10.1093/applin/amu059

Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & F. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21-46). John Benjamins. https://doi.org/10.1075/lllt.32.02bul

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., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing, 26*, 66-79. https://doi.org/10.1016/j.jslw.2014.09.006

Hunt, K. W. (1965). *Grammatical structures written at three grade levels (Research Report No. 3)*. National Council of Teachers of English.

IELTS (n.d.). *IELTS writing band descriptors*. Retrieved July 25, 2024, from https://s3.eu-west-2.amazonaws.com/ielts-web-static/production/Guides/ielts-writing-band-descriptors.pdf

Kang, O., Egbert, J., & Miao, Y. (2023). Speaking and writing features: Distinguishing IELTS proficiency levels and progression over time. *IELTS Research Reports Online Series, No. 1/23*.

Kim, H. (2020). Nominal modifiers in argumentative essays as discriminators for writing course placement decisions. *English Teaching, 75*(3), 3-24.

Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication* [Unpublished doctoral dissertation, Georgia State University]. https://scholarworks.gsu.edu/alesl_diss/35/

Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly, 49*(4), 757-786. http://doi.org/10.1002/tesq.194

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal, 102*, 333-349. https://doi.org/10.1111/modl.12468

Kyle, K., Crossley, S., &Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition, 43*(4), 781-812. https://doi.org/10.1017/S0272263120000546

Lan, G., Lucas, K., & Sun, Y. (2019). Does L2 writing proficiency influence noun phrase complexity? A case analysis of argumentative essays written by Chinese students in a first-year composition course. *System, 85*, 102116. https://doi.org/10.1016/j.system.2019.102116

Li, Y., Lin, S., Liu, Y., & Lu, X. (2023). The predictive powers of fine-grained syntactic complexity indices for letter writing proficiency and their relationship to pragmatic appropriateness. *Assessing Writing, 56*, 100707. https://doi.org/10.1016/j.asw.2023.100707

Li, Y., & Yang, R. (2023). Assessing the writing quality of English research articles based on absolute and relative measures of syntactic complexity. *Assessing Writing, 55*, 100692. https://doi.org/10.1016/j.asw.2022.100692

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474-496. https://doi.org/10.1016/j.asw.2022.100692

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36-62. https://doi.org/10.5054/tq.2011.240859

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing, 34*(4), 493-511. https://doi.org/10.1177/0265532217710675

Müller, A., & Han, W. (2022). IELTS Writing band scores 5.5-7.5: Grammatical error rates, stakeholder perceptions, and risk. *IELTS Research Reports Online Series, No. 1/22*.

Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555-578. https://doi.org/10.1093/applin/amp044

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492-518. https://doi.org/10.1093/applin/24.4.492

Ravid, D., & Berman, R. A. (2010). Developing noun phrase complexity at school age: A text-embedded cross-linguistic analysis. *First Language, 30*, 3-26. https://doi.org/10.1177/0142723709350531

Seo, N., & Oh, S. Y. (2024). Development of clausal and phrasal complexity in L2 writing: A case of argumentative essays of Korean college students. *English for Specific Purposes, 73*, 46-60. https://doi.org/10.1016/j.esp.2023.09.003

Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly, 47*(2), 420-430. https://doi.org/10.1002/tesq.91

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. University of Hawaii Press.

Zhang, X., & Lu, X. (2022). Revisiting the predictive power of traditional vs. fine-grained syntactic complexity indices for L2 writing quality: The case of two genres. *Assessing Writing, 51*, 100597. https://doi.org/10.1016/j.asw.2021.100597

Zhang, Y., & Ouyang, J. (2023). Linguistic complexity as the predictor of EFL independent and integrated writing quality. *Assessing Writing, 56*, 100727. https://doi.org/10.1016/j.asw.2023.100727