



A CORPUS-BASED LEXICO-GRAMMATICAL ANALYSIS OF IELTS ACADEMIC WRITING TASK 1 AND ITS PEDAGOGICAL IMPLICATIONS

Sharipova Zuxra Rustam qizi

O'zbekiston davlat jahon tillari universiteti stajer-o'qituvchisi

<https://doi.org/10.5281/zenodo.21125302>

ARTICLE INFO

Qabul qilindi: 26-iyun 2026 yil
Ma'qullandi: 28-iyun 2026 yil
Nashr qilindi: 30-iyun 2026 yil

KEYWORDS

Writing plays a critical role in high-stakes language assessment, particularly in the International English Language Testing System (IELTS). IELTS Academic Writing Task 1 requires candidates to produce a concise report describing visual data such as graphs, charts, tables, maps, or processes within a limited time frame.

ABSTRACT

Corpus linguistics has become an influential methodological framework in second language acquisition (SLA) research because it enables systematic, empirical analysis of authentic language use. By examining large, principled collections of texts, corpus-based approaches move beyond intuition-based descriptions of language and instead identify recurring lexical and grammatical patterns that characterise particular genres. In SLA, learner corpora are especially valuable, as they provide direct insight into how language learners actually produce written and spoken texts, revealing both emerging competencies and persistent areas of difficulty. As a result, corpus linguistics has increasingly informed research on interlanguage development, assessment, and pedagogical practice.

Corpus linguistics has become an influential methodological framework in second language acquisition (SLA) research because it enables systematic, empirical analysis of authentic language use. By examining large, principled collections of texts, corpus-based approaches move beyond intuition-based descriptions of language and instead identify recurring lexical and grammatical patterns that characterise particular genres. In SLA, learner corpora are especially valuable, as they provide direct insight into how language learners actually produce written and spoken texts, revealing both emerging competencies and persistent areas of difficulty. As a result, corpus linguistics has increasingly informed research on interlanguage development, assessment, and pedagogical practice.

Writing plays a critical role in high-stakes language assessment, particularly in the International English Language Testing System (IELTS). IELTS Academic Writing Task 1 requires candidates to produce a concise report describing visual data such as graphs, charts, tables, maps, or processes within a limited time frame. Although shorter than Task 2, Task 1 contributes substantially to the overall writing band score and is often perceived by learners as deceptively challenging. Successful performance depends on the accurate and appropriate use

of lexico-grammatical resources to summarise data, describe trends, make comparisons, and present an overview, all while maintaining clarity, cohesion, and grammatical accuracy.

Corpus linguistics and language learning

Corpus linguistics is widely defined as the empirical study of language based on large, principled collections of naturally occurring texts stored in electronic form (McEnery & Hardie, 2012; Sinclair, 1991). Unlike intuition-based approaches, corpus linguistics enables researchers to identify recurrent patterns of lexis and grammar that characterise particular genres, registers, or proficiency levels. This empirical orientation has made corpus linguistics especially influential in second language acquisition (SLA) and English for Academic Purposes (EAP), where accurate descriptions of authentic language use are essential.

Within SLA, learner corpora play a central role because they provide direct evidence of how learners actually use language. Learner corpus research has consistently demonstrated that second language writing is shaped by recurrent, non-random patterns, including overuse of high-frequency vocabulary, limited phraseological range, and persistent grammatical difficulties (Granger, 2009). Such findings have contributed to a more nuanced understanding of interlanguage development and have challenged traditional assumptions about learner errors as isolated or unsystematic.

A commonly discussed methodological distinction in corpus linguistics is that between **corpus-based** and **corpus-driven** approaches. Corpus-based studies typically begin with predefined linguistic categories (e.g., articles, tense, cohesive devices) and use corpus data to quantify and refine these categories. In contrast, corpus-driven studies allow patterns to emerge inductively from the data, often through frequency-based procedures such as n-gram or lexical bundle extraction (Tognini-Bonelli, 2001). In practice, many applied studies adopt a hybrid approach, combining corpus-driven discovery with corpus-based interpretation. This mixed orientation is particularly appropriate for highly conventionalised genres such as IELTS Academic Writing Task 1, where both predefined grammatical features and emergent phraseological patterns are pedagogically relevant.

Corpus-based studies on IELTS and EFL writing, with a focus on Task 1

Despite the global prominence of IELTS, corpus-based research on IELTS writing remains relatively limited compared to studies of university academic writing. Among the most relevant contributions is Pearson's (2021) corpus-based comparison of lexical bundles in IELTS Academic Writing Tasks 1 and 2. Using a purpose-built corpus, Pearson demonstrates that Task 1 exhibits a distinctive phraseological profile, characterised by bundles associated with data description, comparison, and summarisation. Importantly, the study highlights that Task 1 writing is more restricted and formulaic than Task 2, reflecting its specific communicative purpose.

Further evidence linking phraseology to IELTS performance is provided by Saadatara et al. (2023), who examined lexical bundle use across IELTS band score levels. Their findings indicate that higher-band scripts employ a wider range of structurally and functionally diverse bundles, whereas lower-band scripts tend to rely on repetitive and semantically vague sequences. By explicitly connecting corpus findings to IELTS band descriptors, this study underscores the relevance of corpus analysis for assessment-oriented pedagogy.

Although not exclusively focused on Task 1, research on stance and evaluative language in IELTS writing offers complementary insights. Işık-Taş (2018) shows that effective IELTS

writing involves careful management of stance and discourse positioning, which has implications for Task 1 in terms of neutrality, cautious generalisation, and appropriate overview statements. Together, these studies suggest that successful IELTS writing depends on mastery of a limited but highly conventionalised set of lexico-grammatical resources.

While existing research confirms the value of corpus approaches for understanding IELTS writing, there remains a need for focused learner-corpus analyses of IELTS Academic Writing Task 1 that integrate lexico-grammatical description with pedagogical application. In particular, few studies combine a clearly defined learner corpus with systematic analysis of lexical, grammatical, and cohesive features and relate these directly to IELTS band descriptors. The present study addresses this gap by offering a corpus-based lexico-grammatical profile of Task 1 learner writing and by demonstrating how these findings can inform targeted instruction.

Teaching Implications

The lexical and grammatical analyses highlight several implications for teaching IELTS Academic Writing Task 1. Overall, the findings suggest that instruction should focus on genre-specific language patterns rather than on general vocabulary development or isolated grammar rules. Both analyses show that Task 1 writing relies on a limited but highly functional set of recurring lexical items and grammatical structures.

From a lexical perspective, learners need focused instruction in key semantic domains such as quantification, change and trends, comparison, visual data description, and time reference. Teaching vocabulary in semantic groups can help learners understand how words function together to describe data. The frequent use of items such as *number*, *percentage*, *increase*, and *overall* also indicates that accuracy and appropriate use are more important than lexical variety in this genre. Teachers should therefore prioritize correct and conventional usage rather than encouraging unnecessary synonym variation.

This study explored the linguistic characteristics of IELTS Academic Writing Task 1 through a corpus-based approach and examined how corpus findings can inform English language teaching. The lexical analysis showed that Task 1 writing relies on a restricted vocabulary centred on quantification, trends, comparison, and time reference. The grammatical analysis revealed recurrent structures that function as building blocks of the genre, highlighting the close relationship between lexis and grammar in academic data description.

References:

1. Biber, D., Conrad, S. and Reppen, R. (1998) *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
2. Farr, F. and Leńko-Szymańska, A. (2024) 'Corpus linguistics and language pedagogy: bridging research and practice', *Applied Linguistics*, 45(1), pp. 1–18.
3. Granger, S. (2015) 'The contribution of learner corpora to second language acquisition and foreign language teaching', *ReCALL*, 27(1), pp. 1–19.
4. Hunston, S. (2002) *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
5. Hyland, K. (2008) *Genre and academic writing in the disciplines*. London: Continuum.
6. McEnery, T. and Hardie, A. (2012) *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.

7. Ngo, T. and Chen, H. (2024) 'Corpus-based approaches to ESL/EFL writing development', *Journal of English for Academic Purposes*, 67, pp. 1–14.



**INNOVATIVE
ACADEMY**