

SYNTACTIC ERROR PATTERNS IN VIETNAMESE UNIVERSITY ENGLISH AS A FOREIGN LANGUAGE WRITING: A DIAGNOSTIC INTERPRETATION USING A HYBRID TAXONOMY

Vo Ngoc Anh Minh¹

Diep Vu Dang Quang² Nguyen Truong Giang³

^{1,2}International University, Vietnam National University Ho Chi Minh City; ³Thanh Do University

Email: vongocanhminhcontactwork@gmail.com¹; dangquangdiep0510@gmail.com²;

ntgiang@thanhdouni.edu.vn³

Received: 01/11/2025; Reviewed: 4/12/2025; Revised: 8/12/2025; Accepted: 25/12/2025

DOI: <https://doi.org/10.58902/tcnckhpt.v4i4.311>

Abstract: *Many university English as a Foreign Language learners continue to experience persistent syntactic difficulties in academic writing despite extensive exposure to formal grammar instruction. This study examines these challenges among first-year English majors at a Vietnamese university, a cohort working at the B2–C1 threshold of academic literacy. Using a mixed-methods design, 49 problem–solution essays were analyzed through a hybrid taxonomy that integrates the structural precision of Dulay et al.’s (1982) Surface Strategy Taxonomy with the pedagogical distinctions proposed by Ferris (2005). The analysis identified a marked predominance of Misformation errors (48.36%), far exceeding Omission (19.07%) and other types. Interpreted alongside survey data indicating a strongly rule-focused instructional history, these patterns suggest that the difficulties observed relate not to the absence of grammatical knowledge but to the demands of applying it consistently in extended written production. Article omissions, although consistent with L1-related trends documented in previous research, co-occurred with structurally ambitious attempts, pointing to an interaction of linguistic background, instructional experience, and task conditions. The findings highlight the value of diagnostic, context-sensitive error analysis over descriptive counts alone and suggest pedagogical attention to meaningful writing practice and reflective engagement with form–function relations.*

Keywords: *Rule-based instruction; Syntactic errors; Misformation; Interlanguage; Diagnostic error analysis; Academic writing.*

1. Introduction

The strategic integration of Vietnam into the global economy, enshrined in national policy frameworks like Decision 2080/QĐ-TTg, has positioned English proficiency as a critical asset for educational and professional advancement. This has precipitated a significant shift in English language teaching (ELT), particularly in higher education, where the focus has intensified on developing advanced communicative skills, with academic writing being a cornerstone of academic success and future employability (Mai & Ly, 2024). However, within this English as a Foreign Language (EFL) context, a persistent paradox emerges: despite years of formal instruction and increased exposure, many Vietnamese learners, even English majors, continue to struggle with syntactic accuracy in

their written production (Khang et al., 2021; Chi, 2020).

This challenge is theoretically significant. Writing, as a cognitively demanding generative skill, requires the seamless integration of complex syntactic knowledge, making errors an inevitable feature of the second language (L2) acquisition process (Bello, 2017). Rather than mere failures, errors are now widely understood through the lens of Error Analysis (EA) as invaluable windows into the learner’s interlanguage—the systematic, evolving linguistic system they construct (Selinker, 1972; Corder, 1967). For advanced learners, error patterns are particularly revealing; they can indicate areas of fossilization, the overextension of learned rules, or the struggle to manage complexity under real-time production pressures (Han, 2004; He &

Oltra-Massuet, 2023). Therefore, analyzing errors moves beyond a deficit model to become a crucial diagnostic tool for understanding L2 development and refining pedagogical interventions (Ellis & Barkhuizen, 2005).

A review of recent Vietnamese scholarship indicates that studies employing Error Analysis have predominantly adopted a descriptive orientation, with research efforts largely directed toward cataloguing surface-level grammatical or syntactic deviations. Within this line of work, Nhu and Ho (2024) examine recurrent syntactic errors in argumentative essays and provide detailed frequency-based accounts of students' difficulties. Dinh (2023) similarly documents the distribution of grammatical errors among pre-intermediate learners with the purpose of informing remedial instruction. Dung and Tran (2023), though situated in translation studies, also rely on EA to identify error patterns and their likely sources. These studies contribute valuable empirical baselines; however, they share a tendency to prioritize taxonomic identification over inquiries into how such errors relate to learners' developing syntactic competence or their procedural command of academic writing conventions. What remains less explored is the extent to which syntactic difficulty in authentic academic writing reflects deeper challenges in consolidating rule-based knowledge into fluent written performance, despite prior exposure to explicit grammar instruction. This under-addressed dimension suggests an opportunity to extend Vietnamese EA research beyond descriptive profiling toward an account that attends to learners' underlying syntactic decision-making and the cognitive processes shaping persistent error patterns.

To address these limitations, this study employs an explanatory sequential mixed-methods design to investigate the syntactic competence of English-majored undergraduates at a Vietnamese international university. The research is guided by the following questions:

1. *What is the frequency and typology of syntactic errors in the academic writing of Vietnamese English-majored undergraduates at the B2-C1 proficiency level?*
2. *How do these learners' reported English*

learning experiences provide a diagnostic lens for interpreting the predominant error patterns?

To answer these questions, this research applies a hybrid analytical framework integrating the linguistic precision of Dulay, Burt, and Krashen's (1982) Surface Strategy Taxonomy with the pedagogical orientation of Ferris's (2005) taxonomy. By analyzing problem-solution essays and triangulating the findings with data on learners' instructional histories, this study aims to move beyond superficial error categorization. The goal is to contribute a diagnostic, contextually grounded understanding of Vietnamese university EFL learners' interlanguage at the B2-C1 threshold, offering insights that can inform the development of more responsive writing pedagogies aligned with the nation's strategic objectives for English education.

2. Research overview

This review constructs a critical pathway from foundational theoretical principles to an integrated conceptual framework for diagnosing syntactic errors in the academic writing of Vietnamese university EFL learners at the B2-C1 threshold. It begins by establishing Error Analysis (EA) as the essential theoretical lens, then critically examines the taxonomic tools for categorizing errors, synthesizes empirical findings to reveal a central gap, and culminates in the presentation of a hybrid diagnostic framework that operationalizes this pathway from theory to pedagogical insight.

2.1. Error Analysis: From Theoretical Foundation to Diagnostic Tool

The systematic study of learner errors is fundamentally grounded in the paradigm of Error Analysis (EA), which emerged as a transformative response to the limitations of Contrastive Analysis. EA reconceptualized errors from being signs of failure to being inevitable, systematic indicators of a learner's developing interlanguage - a legitimate, rule-governed linguistic system (Corder, 1967; Selinker, 1972). This pivotal shift acknowledged that errors originate from multiple sources, including both interlingual transfer (L1 influence) and intralingual processes like overgeneralization of L2 rules (Richards, 1970). This theoretical foundation positions the learner as an active hypothesis-tester, making errors a valuable

source of insight into the acquisition process itself.

The diagnostic potential inherent in EA is profound: by meticulously analyzing errors, researchers and educators can gain a window into learners' underlying competence, thereby tailoring instruction to address developmental needs (Ellis, 1997). This potential is especially critical at advanced proficiency levels, where syntactic inaccuracies risk becoming fossilized, persistently hindering academic expression (Han, 2004).

However, a critical appraisal of EA's application, particularly in EFL contexts like Vietnam, reveals a significant chasm between its theoretical promise and its practical execution. While studies consistently employ EA to catalogue frequent error types (e.g., Dinh, 2023; Hidayat et al., 2020), the approach frequently remains descriptive rather than diagnostic. The common practice of stopping at frequency counts constitutes a significant limitation; it identifies what errors occur but seldom probes the conditions under which they persist, even among learners who have received explicit instruction (Khatter, 2019). This descriptive tendency ultimately restricts EA's utility in shaping long-term, effective pedagogical interventions. The theoretical bedrock is solid, but the methodological superstructure often lacks the explanatory power needed to bridge theory and practice.

To respond to these limitations while maintaining EA's diagnostic potential, the present study adopts a data selection strategy that is explicitly aligned with the investigation of learners' syntactic performance in academic discourse. In selecting data for analysis, the study foregrounds coherence between the nature of learner production and the analytical aims. Given that the purpose is to examine how undergraduate learners deploy syntactic resources in academic writing, the corpus consists of intact argumentative essays produced within a shared instructional context. The focus on a single written genre and a single learner cohort is intentional: it enables a concentrated examination of syntactic behavior within a controlled communicative task, rather than dispersing analytical attention across modality- or genre-

driven variation that lies beyond the study's scope. This rationale ensures internal alignment between research questions, data type, and analytical focus, thereby supporting a more precise interpretation of learners' syntactic performance.

2.2. Taxonomic Frameworks: Linguistic Precision Versus Pedagogical Utility

To operationalize EA, researchers rely on taxonomic frameworks for classifying errors. The literature reveals two dominant traditions, each with distinct strengths and limitations that shape the analytical outcome.

The first tradition, exemplified by the Surface Strategy Taxonomy (Dulay, Burt, & Krashen, 1982), offers a robust, linguistically precise system for categorization. It classifies errors based on their structural manifestation, Omission, Addition, Misformation, and Misordering, providing granular, descriptive clarity about how a learner's output deviates from the target form. This taxonomy is exceptionally powerful for identifying the specific structural nature of an error.

In contrast, the second tradition, championed by Ferris (2005), prioritizes pedagogical utility. Her taxonomy categorizes errors (e.g., fragments, run-ons) based on their amenability to correction via rule-based instruction, distinguishing between "treatable" and "untreatable" errors. This framework is highly practical for classroom teachers, as it directly informs instructional focus. However, it can lack the linguistic granularity required for a deep analysis of the underlying interlanguage system.

The critical pathway forward does not lie in choosing one taxonomy over the other, but in recognizing their complementary nature. Dulay et al.'s (1982) framework provides the diagnostic precision, while Ferris's (2005) offers the pedagogical relevance. A truly comprehensive analysis requires a synthesis that leverages the strengths of both.

Critical Synthesis of Empirical Research: Highlighting the Diagnostic Void

Empirical studies employing these and other taxonomies have yielded valuable, yet often inconsistent, findings. Some research identifies Omission as the most frequent error (Nadya &

Muthalib, 2021; Pradnyaswari et al., 2022), while other studies find Misformation to be predominant, particularly at intermediate-to-advanced levels (Ma'mun, 2016; Setiaji et al., 2023). This inconsistency suggests that error patterns are influenced by variables such as learner proficiency, instructional context, and writing genre-factors that are frequently underexplored in purely descriptive studies.

More critically, a recurrent pattern across EFL scholarship-and clearly visible in the Vietnamese literature-is the limited use of EA for explanatory purposes. Studies such as Dung and Tran (2023) and Nhu and Ho (2024) offer detailed accounts of error types, yet seldom investigate the underlying conditions that give rise to these persistent patterns, particularly among learners who have experienced sustained exposure to explicit instruction. This tendency mirrors trends observed in many EFL settings, where descriptive taxonomies dominate and diagnostic insight remains underdeveloped. The need therefore arises for an analytical framework capable of integrating detailed taxonomic precision with explanatory depth, enabling a more coherent understanding of the sources of syntactic errors.

2.3. Conceptual Framework: A Hybrid Diagnostic Pathway for Pedagogical Insight

The present study adopts a hybrid conceptual framework that traces a coherent analytical trajectory from the identification of syntactic deviations to the diagnostic interpretation of the conditions under which they arise. This trajectory responds to long-standing critiques of Error Analysis, especially the observation that descriptive accounts often end at taxonomic listing without clarifying how structural patterns interact with instructional experiences or task demands. Within this framework, the analysis unfolds through an interlinked sequence in which error typology, error frequency, and diagnostic interpretation function as mutually dependent layers rather than discrete methodological units.

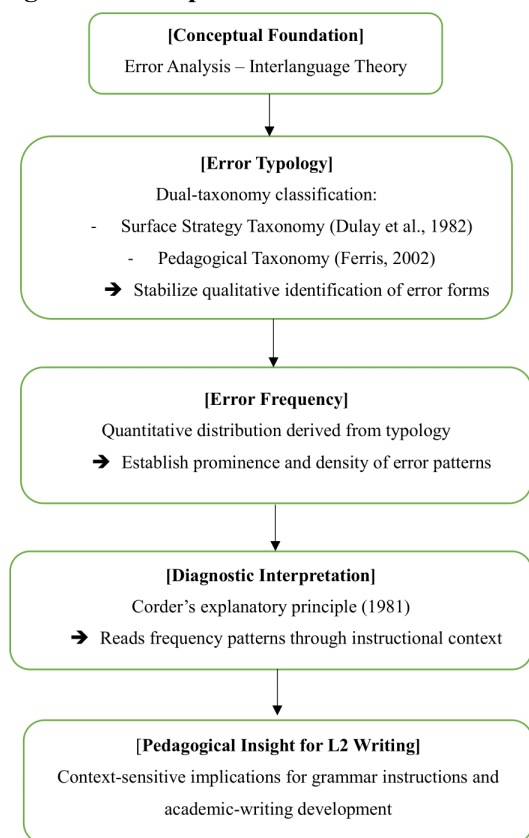
The first layer, error typology, is constructed through the combined application of Dulay, Burt, and Krashen's Surface Strategy Taxonomy and Ferris's pedagogical taxonomy. This dual system allows for the principled identification of

syntactic deviations in terms of both their structural configuration and their instructional relevance. It therefore establishes the qualitative foundation upon which all subsequent analytical steps depend.

Once typological categories have been stabilized, the analysis proceeds to the second layer, error frequency patterns. Frequency is conceptualized not as an autonomous quantitative outcome but as an empirical distribution that acquires meaning only when grounded in theoretically motivated classification. This placement is consistent with foundational EA claims that numerical prominence must be interpreted in light of structured error categories. Frequency thus serves as an intermediary layer that reveals the density and salience of particular syntactic tendencies and prepares the ground for interpretive analysis.

The third layer, diagnostic interpretation, draws on Corder's notion that explanation requires situating learner output within the conditions that shape performance. In this study, such interpretation is anchored in learners' reported instructional histories and the demands of academic writing tasks, rather than in speculative psycholinguistic inference. Reading the frequency patterns-particularly the predominance of Misformation errors-through this contextual lens allows the analysis to illuminate how learners negotiate explicit, rule-based knowledge under the pressures of extended written production.

Together, these layers form a sequential yet interdependent mechanism in which structural classification generates meaningful frequency patterns, and those patterns acquire explanatory value only when contextualized within learners' educational trajectories. By mapping the movement from typology to frequency to diagnostic reasoning, the framework advances a form of Error Analysis that retains descriptive rigor while extending its interpretive capacity in a controlled, theory-informed manner. It thereby provides a transparent rationale for how the study moves from surface deviations toward context-sensitive pedagogical insight, without collapsing analytic stages or overstating causal claims.

Figure 1. Conceptual Framework of the Study

3. Research methods

3.1. Research Design

This study employed an explanatory sequential mixed-methods design (Creswell & Plano Clark, 2017), with a primary quantitative phase followed by a supplementary qualitative phase. This design was selected to first identify and quantify the most frequent syntactic errors in learners' essays (quantitative phase) and then use qualitative data to explore the potential learning experiences and strategies that may underlie these error patterns (qualitative phase). The integration of methods aimed to provide a more nuanced, diagnostic understanding of the syntactic challenges faced by the participants, moving beyond mere description toward explanatory insight.

3.2. Participants and Context

The participants were 49 first-year students majoring in English at International University, Vietnam National University Ho Chi Minh City. All participants were enrolled in a mandatory "Writing 1" course, designed for learners at the B2-C1 level according to the Common European

Framework of Reference (CEFR). Entry into this course required a minimum IELTS score of 6.0 or a passing grade on the university's internal placement test, ensuring a relatively homogeneous level of upper-intermediate to advanced proficiency.

3.3. Data Collection Instruments and Procedures

Data collection occurred in two sequential phases over a one-month period. Phase 1 focused on capturing learners' spontaneous syntactic performance under uniform academic-writing conditions. All participants completed an in-class problem–solution essay in week 10, choosing one of three pre-set topics to limit content-driven variation. They wrote a 300-word essay within 60 minutes ($\pm 15\%$; Weigle, 2002). Students had previously received instruction on common writing issues and on grammatical structures relevant to the analysis, which ensured that the deviations observed were not attributable to unfamiliarity with target forms.

Phase 2 involved an online background survey administered after the error analysis. Adapted from Ellis's learner-profile framework, the instrument elicited information on learners' L1, duration of English study, instructional environments, and retrospective accounts of grammar and writing instruction. These data were used solely to contextualize the error patterns identified in Phase 1, providing the evidentiary basis for the diagnostic interpretation without altering the original coding.

3.3. Data Analysis

The data were analyzed in alignment with the sequential mixed-methods design, beginning with a quantitative examination of syntactic errors followed by a qualitative interpretation of the survey responses.

Quantitative Analysis of Syntactic Errors

Syntactic errors were identified and categorized through a descriptive Error Analysis procedure (Ellis, 1997). Each essay was reviewed line-by-line and coded using a hybrid taxonomy that combined the structural categories of the Surface Strategy Taxonomy (Dulay et al., 1982) with the instructional tractability distinctions proposed by Ferris (2005). All coded instances

were entered into a spreadsheet and summarized using descriptive statistics (raw counts and proportional frequencies) to map the most recurrent syntactic patterns in the corpus.

To support consistency, two trained coders independently coded a subset of essays using a shared coding guide. Inter-coder agreement, calculated as percent agreement, reached approximately 90%, after which discrepancies were discussed until a unified interpretation was obtained.

Qualitative Analysis of Survey Responses

Open-ended survey responses were examined using reflexive thematic analysis (Braun & Clarke, 2006). The procedure involved familiarization with the responses, inductive coding, and the development of themes that captured recurrent features of learners' instructional experiences. The thematic patterns served to illuminate contextual conditions relevant to the syntactic tendencies observed in the quantitative strand.

Integration of Strands

This integrative step aligns with the

Table 1. Frequency and Distribution of Syntactic Errors (N=49 essays)

| Error Type | Frequency | Percentage (%) | Errors/100 words | Rank |
|--------------|-------------|----------------|------------------|------|
| Misformation | 487 | 48.36 | 2.97 | 1 |
| Omission | 192 | 19.07 | 1.17 | 2 |
| Run-ons | 111 | 11.02 | 0.68 | 3 |
| Addition | 105 | 10.43 | 0.64 | 4 |
| Fragments | 84 | 8.34 | 0.51 | 5 |
| Misordering | 28 | 2.78 | 0.17 | 6 |
| Total | 1007 | 100 | 6.14 | |

As Table 1 demonstrates, **Misformation** errors were overwhelmingly dominant, accounting for nearly half of all deviations (48.36%). A deeper analysis of these 487 errors revealed that the most prevalent sub-categories were:

- **Verb Form and Complementation Errors (~40% of Misformations):** These included incorrect verb patterns (e.g., *suggest to do* instead of *suggest doing*), and errors in tense/aspect sequencing.

- **Article Misuse (~25% of Misformations):** This involved the incorrect selection of definite/indefinite articles in context, going beyond simple omission.

explanatory purpose of the design outlined in Section 3.1. Interpretation occurred through connecting the descriptive error patterns with the thematic insights about instructional exposure. This integrative reading enabled a diagnostic account grounded in the dataset while avoiding claims beyond what the evidence could substantiate.

4. Research results

This section presents the findings of the study, structured to first address the primary research question on the frequency and types of syntactic errors, and then to integrate the qualitative data on learners' backgrounds to provide initial, data-grounded insights into potential sources of these errors.

4.1. Dominance of Misformation Errors in L2 Writing at the B2-C1 Threshold

Analysis of the 49 essays (total words: 16,424) revealed a total of 1,007 syntactic errors. As detailed in Table 1, the distribution of errors was not even, with a clear dominance of one specific type.

- **Preposition Errors (~20% of Misformations):** Errors in the choice of prepositions following verbs, adjectives, or nouns.

A representative example is: "*The decision request people spend almost time in their life to find a way to learning effective.*" (Code 21). This sentence exhibits multiple Misformation errors: the choice of infinitive "*to find*" after "*spend time*" (where a gerund "*finding*" is required), the erroneous gerund "*learning*" after "*to*" (where the infinitive "*learn*" is correct), and the adjectival form "*effective*" used where the adverbial form "*effectively*" is needed to modify the verb.

The second most common error type was Omission (19.07%), primarily of articles,

prepositions, and relative pronouns, as in: “*Sports are the most powerful thing () can reduce stress.*” (Code 36). The relative pronoun “*that*” or “*which*” is absent, creating a syntactic fragment. Run-ons (11.02%) and Additions (10.43%) were also frequent, while Fragments (8.34%) and Misordering (2.78%) were comparatively rare.

4.2. Learner Backgrounds: A Profile of Rule-Based Instruction

The survey data from 40 participants provides a crucial diagnostic lens through which to interpret the quantitative error patterns. This data paints a coherent and revealing picture of a homogenous learner cohort shaped by a specific, rule-dominant instructional history, which directly informs the prevalence and nature of their syntactic challenges.

4.2.1. A Homogenous Cohort with Extensive Exposure

Figure 2. First language



The participant profile establishes a controlled context for analysis. As illustrated in Figure 1, the cohort is overwhelmingly homogenous in L1, with 97.5% (n=39) of participants identifying Vietnamese as their first language. This homogeneity effectively minimizes cross-linguistic variation as a confounding variable, allowing for a sharper focus on intralingual and instructional factors as primary sources of syntactic errors. Furthermore, Figure 2 details the participants' extensive experience with English, with learning durations ranging from 4 to 15 years. A significant majority (65%, n=26) fall within the 10- to 13-year range, indicating a cohort with substantial cumulative exposure to the language. This extended learning period is critical, as it suggests that the syntactic errors observed are not merely a function of limited exposure but are characteristic of challenges that persist into advanced stages of acquisition.

4.2.2. A Predominantly Rule-Based Instructional Ecosystem

Figure 3. Years of Learning

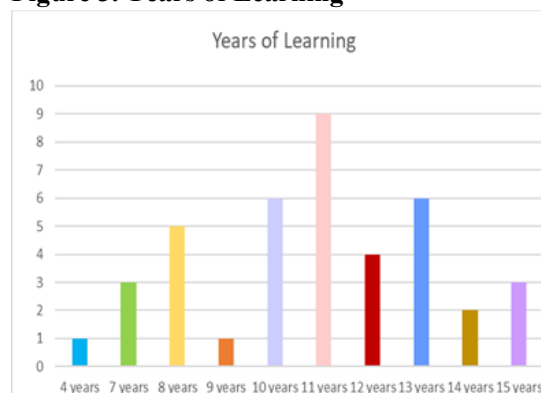


Figure 4. English grammar and students' learning environment

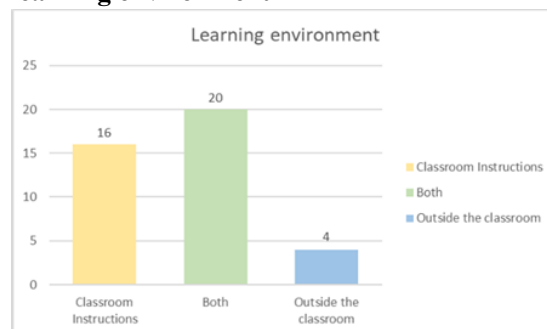
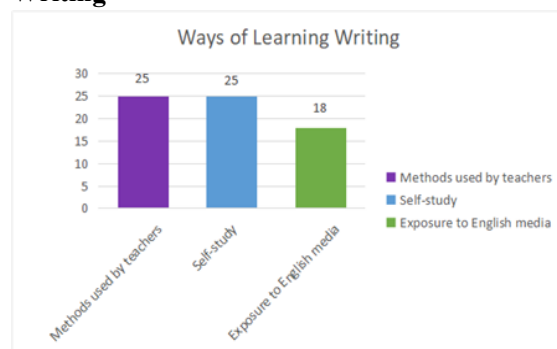


Figure 5. Approaches to learn English Writing



The most significant findings pertain to the instructional environment that has shaped these learners. The data reveals a consistent emphasis on formal, explicit instruction across multiple dimensions. Figure 3 shows that the primary learning environment is either exclusively the classroom (n=16) or a combination of classroom and external sources (n=20), with purely external learning being rare (n=4). This indicates that the learners' linguistic development has occurred within a structured, pedagogical context.

This focus on formal instruction is further

clarified in Figures 4 and 5. Figure 4, detailing the "Ways of Learning Writing," demonstrates that teacher-led methods and self-study are the dominant, equally utilized strategies (n=25 each), significantly outweighing learning through media exposure (n=18). This suggests a learning culture oriented towards deliberate practice and explicit knowledge acquisition.

Most critically, Figure 5 provides unambiguous evidence of a rule-based pedagogical paradigm. When asked how grammar was previously taught, a vast majority of respondents (85%, n=34) identified "Rule-based lessons" as the primary method. This proportion is substantially higher than those highlighting "Examples" (n=18) or, most tellingly, "Writing practice" (n=24). The disparity between the high incidence of rule-based instruction and the comparatively lower emphasis on writing practice is particularly revealing. It suggests that these learners have been equipped with a strong foundation in declarative knowledge (knowing the rules) but have had comparatively fewer opportunities to proceduralize this knowledge through extensive, meaningful production practice.

Figure 6. Approaches to teach grammar



Figure 6 quantitatively illustrates the predominant methods through which grammar was taught in the learners' previous educational experiences, as retrospectively reported by the participants.

The data reveals a stark pedagogical landscape. The method "Rule-based lessons" is overwhelmingly dominant, with a frequency of n=34 (85%). This signifies that the vast majority of learners were exposed to grammar instruction that prioritized the explicit explanation of grammatical rules, likely involving metalinguistic terminology and deductive teaching methods. This finding is not merely descriptive; it is diagnostic. It

establishes that these learners entered the study with a highly developed foundation in declarative knowledge—they had been explicitly taught *what* the rules of English grammar are.

The other two methods, "Writing practice" (n=24) and "Examples" (n=18), are significantly less frequent. This disparity is highly significant for interpreting the error data. The fact that "Writing practice" is reported less frequently than "Rule-based lessons" suggests a potential imbalance in their instructional history. While learners received extensive information about the language, they may have had comparatively fewer opportunities to proceduralize this knowledge through sustained, meaningful production in writing tasks. Similarly, the lower frequency of learning through "Examples" may indicate a pedagogical approach that favored direct rule presentation over inductive, exemplar-based learning.

4.3. Toward an Integrated Analysis: Linking Error Patterns to Learning Experiences

The juxtaposition of the quantitative error data and the qualitative background data suggests a critical relationship. The high frequency of Misformation errors—which involve selecting an incorrect form despite knowing a rule is required—among learners from heavily rule-based instructional backgrounds points to a potential gap between declarative knowledge (knowing the rule) and procedural knowledge (applying it accurately in complex writing).

This is further supported by a qualitative cross-analysis of individual responses and error types. For instance, learners who described their instruction as "memorizing grammar rules" frequently produced errors of overgeneralization, such as the erroneous application of the "to + gerund" pattern in places where the infinitive was required (e.g., *to learning*). This pattern suggests that these learners possess explicit metalinguistic knowledge but struggle with its contextual application, leading to what could be termed "rule-based misapplications."

Conversely, errors of Omission (e.g., article omission), which are often strongly linked to L1 transfer, were present across the board, but did not show a clear correlation with any specific instructional background, indicating their

persistence as a distinct challenge likely rooted in the fundamental differences between Vietnamese and English.

These integrated, data-driven insights suggest that the error profile of this cohort is not random but is systematically shaped by their instructional experiences. The dominance of Misformation errors appears to be a hallmark of learners who have been equipped with declarative rules but have had insufficient opportunity to proceduralize that knowledge through meaningful, contextualized practice. This sets the stage for a discussion that can further theorize this disconnect between knowledge and application.

5. Discussion

The aim of this study was to revisit the diagnostic potential of Error Analysis within the Vietnamese EFL context, where prior research has largely remained descriptive and has less frequently examined the instructional conditions shaping learners' syntactic performance. The findings of this study respond to this gap by illustrating how a hybrid taxonomy, when interpreted through a diagnostic orientation anchored in instructional context, can illuminate the underlying tendencies that generate recurrent syntactic patterns. This section integrates the quantitative and qualitative strands and situates the findings within the scholarly conversations articulated in the literature review.

5.1. Revisiting the Hybrid Taxonomy: What the Error Patterns reveal

The dominant presence of Misformation errors in the corpus highlights the value of combining the Surface Strategy Taxonomy with Ferris's pedagogical categories. Nearly half of all coded errors (48.36%) fell into this category, far exceeding Omission (19.07%) and all other types. While previous Vietnamese studies have catalogued errors primarily in terms of structural frequency, this hybrid approach provided a more calibrated lens through which to view the learners' syntactic behavior. The high proportion of Misformation suggests not random inaccuracies but patterned misapplications of rules that learners have already encountered. Rather than interpreting this as a sign of linguistic deficiency, the taxonomy allowed these errors to be understood as part of learners' ongoing

engagement with structurally demanding academic writing tasks.

This interpretation is consistent with long-standing interlanguage perspectives reviewed earlier: errors may reflect the learner's attempt to mobilize existing knowledge in increasingly complex contexts, rather than simple gaps in knowledge. The relatively lower incidence of Omission errors (19.07%), particularly when contrasted with the density of misapplied forms, reinforces this developmental reading. The taxonomy therefore provided a structured view of what patterns were present, and it situated those patterns within a more developmentally sensitive perspective than a purely descriptive count could achieve.

5.2. Interpreting the Patterns through Instructional Context

The diagnostic orientation embedded in the framework required that the structural findings be read alongside the survey data. The background responses pointed to extended exposure to rule-focused instruction, a characteristic documented in both local and international EFL research. Notably, 85% of participants reported learning grammar primarily through rule-based lessons, and fewer than half indicated consistent writing practice. When interpreted through Corder's explanatory step, this instructional history offers a grounded way to understand the predominance of Misformation errors. Rather than proposing new cognitive mechanisms or task-induced causal chains, the findings can be understood as the interaction between explicit knowledge accumulated through rule-based instruction and the demands of academic writing tasks that require the rapid integration of that knowledge.

From this perspective, the errors do not signify a failure of instruction but reveal the limits of instruction that prioritizes explicit rule learning over opportunities for practice that consolidates control. This interpretation remains within the evidentiary scope of the study: it does not assume unobserved cognitive processes but relies on the instructional histories that participants themselves reported. In doing so, it moves beyond the descriptive tendency noted in the literature review and demonstrates how EA can support a more situated understanding of why

particular syntactic difficulties recur.

The persistence of article omission, which has been widely attributed to L1 transfer in prior studies, was also observable in the corpus. Although article-related errors constituted roughly a quarter of all Misformations, they did not dominate the overall error profile. Yet the diagnostic orientation encouraged a more nuanced reading: rather than attributing these omissions solely to L1 influence, their co-occurrence with structurally ambitious attempts suggests that learners may have been allocating attention to the syntactic organization required by the academic genre. This does not constitute a causal explanation but reflects the interaction between linguistic background, instructional experience, and task conditions that become visible when error patterns are read diagnostically rather than categorically.

5.3. Interpreting Interlanguage Development without Overextension

The findings invite reflection on how interlanguage development might be conceptualized for learners working at the B2–C1 threshold. The prominence of Misformation errors indicates that these learners are not avoiding syntactic complexity; rather, they are actively engaging with it, sometimes overextending forms they partially control. The distribution of errors - nearly 500 Misformations produced across 49 essays - illustrates this willingness to attempt syntactically demanding structures even when control is uneven. This aligns with interlanguage research that positions errors as evidence of developmental risk-taking. Instead of positing a new theoretical construct, the patterns observed here can be viewed as an instance of how explicit knowledge, when still stabilizing, may manifest as systematic inaccuracies in academically demanding tasks.

Within this interpretation, proceduralization is presented as a well-established concept in SLA that provides a vocabulary to articulate the gap between declarative familiarity with rules and the fluent use of those rules in performance. The findings of this study resonate with earlier arguments in the literature that learners at this proficiency range may display high analytical awareness but uneven deployment of that

knowledge when managing academic writing tasks.

Importantly, this interpretation does not generalize beyond the data: it remains bounded by the instructional histories, genre demands, and analytical framework within which the study was conducted.

5.4. Pedagogical Insights Grounded in Diagnostic Reading

The pedagogical implications that emerge from this diagnostic reading echo the literature's call to bridge the gap between error identification and instructional response. Because the hybrid taxonomy enabled clearer insight into both the structural and instructional dimensions of the errors, particularly the pattern in which Misformations constituted almost half of all deviations and were more than double the frequency of Additions, Run-ons, or Fragments combined, the findings suggest that instruction may benefit from practices that strengthen learners' control over forms they already partially command. This interpretation aligns with the broader literature that highlights the need to support learners in consolidating the link between explicit knowledge and accurate performance in writing.

From this standpoint, genre-aware writing tasks and opportunities for structured output that require learners to activate target syntactic resources within meaningful contexts may be useful in addressing the specific patterns observed. Similarly, feedback practices that prompt attention to form-function relations may complement this process by prompting reflection on form–function relationships. While the present study did not examine feedback practices directly, the diagnostic tendencies observed in the data resonate with pedagogical positions that emphasize metalinguistic engagement as a means of strengthening control.

These implications are not positioned as generalizable prescriptions for all Vietnamese EFL learners, but as context-specific recommendations derived from the analytic configuration of this study. They should therefore be interpreted as possibilities suggested by the diagnostic patterns rather than definitive instructional solutions.

5.5 Reflexive Acknowledgement of Analytical Shaping

Finally, it is important to recognize that the diagnostic insights produced here are shaped by the methodological decisions embedded in the framework. The error patterns reflect, in part, the rhetorical demands of the problem-solution essay, and the hybrid taxonomy foregrounds particular dimensions of learner performance while backgrounding others. This reflexivity aligns with critical perspectives in EA that emphasize the constructed nature of diagnostic interpretations. Acknowledging these conditions does not weaken the findings; instead, it positions them responsibly within their analytical and contextual boundaries.

This discussion has demonstrated how the hybrid diagnostic framework, grounded in established EA scholarship and contextualized through instructional factors, can enrich the interpretation of syntactic error patterns. The disproportionate share of Misformations (48.36%) provides particularly strong evidence for the need to move beyond surface description. By aligning structural classification with contextual understanding, the study moves beyond description toward a more situated, evidence-driven account of learners' syntactic development, one that neither overextends theoretical claims nor detaches itself from the literature that motivates the inquiry.

6. Conclusion

This study examined the syntactic patterns in the academic writing of Vietnamese English majors through a diagnostic Error Analysis framework that combined a structural taxonomy with an instructional lens. The analysis revealed a marked dominance of Misformation errors-nearly

half of all deviations-which, when read alongside learners' heavily rule-based instructional histories, suggests that the central challenge lies less in the absence of grammatical knowledge than in managing that knowledge under the demands of academic writing.

Interpreted through established interlanguage perspectives, these patterns indicate a developing system in which structurally ambitious attempts and persistent omissions coexist. Rather than pointing to a single source, the diagnostic orientation highlighted how instructional experiences, genre expectations, and partial control over grammatical forms intersect in shaping learners' choices. This integrated reading moves beyond simple error listing and offers a more situated understanding of why particular difficulties remain salient at the B2-C1 threshold.

The pedagogical implications that follow are necessarily modest and context-bound. The findings suggest that instructional support may benefit from increased opportunities for meaningful written production and reflective attention to the relationship between form and rhetorical purpose-areas that align with learners' demonstrated needs without generalizing beyond the present dataset.

The study's scope is limited to one genre and a relatively homogenous cohort, and future research could extend the diagnostic approach across tasks or explore the cognitive pressures underlying recurrent Misformations. Within these boundaries, the study provides a focused account of how Vietnamese learners navigate syntactic complexity in academic writing and illustrates the value of diagnostic perspectives that treat errors as evidence of development rather than as isolated deficiencies.

References

- Bello, P. M. (2017). An analysis of grammatical errors in written text by the first year students of Vocational of Economics and Commerce Becora, Dili in school year 2016/2017. *International Journal of Advance Research*, 5(9), 512–520.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
- Chi, P. K. (2020). Error analysis on English compositions and paragraphs on Vietnamese students. *TNU Journal of Science and Technology*, 225(11): 55 – 63. <https://doi.org/10.34238/tnu-jst.3183>.
- Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching*, 5(4), 161–

- 170.
- Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage Publications.
- Dinh, L. T. T. (2023). Grammatical error analysis of EFL earners' English Writing Samples: The case of Vietnamese pre-intermediate students. *International Journal of TESOL & Education*, 3(3), 41–50. <https://doi.org/10.54855/ijte.23341>.
- Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. Oxford University Press. <https://archive.org/details/language20000dul/page/n1/mode/1up>.
- Dung, D. T. M., & Tran, N. T. (2023). An investigation into common errors in Vietnamese-English translation made by college freshmen: A case study at Foreign Trade University – Ho Chi Minh Campus. *Thu Dau Mot University Journal of Science*, 5(4), 482–501. <https://doi.org/10.37550/tdmu.EJS/2023.04.499>.
- Ellis, R. (1997). *Second language acquisition*. Oxford University Press. https://drive.uqu.edu.sa/_/malehyani/files/Rod-Ellis-The-Study-of-Second-Language-A.pdf.
- Ellis, R., & Barkhuizen, G. P. (2005). *Analysing Learner Language*. Oxford University Press.
- Ferris, D. R. (2005). *Treatment of error in second language student writing*. University of Michigan Press.
- Han, Z. (2004). *Fossilization in adult second language acquisition* (Vol. 5). Multilingual Matters.
- He, Q., & Oltra-Massuet, I. (2023). An experimental study on grammatical sensitivity and production competence in Chinese and Spanish EFL learners and its implications on EFL teaching methods. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1096875>.
- Hidayat, D. N., Fitriyani, N., Alek, A., Eviyuliwati, I., & Septiawan, Y. (2020). An investigation into the grammatical errors of students' writing. *Eduvelop: Journal of English Education and Development*, 4(1), 9–16.
- Khang, N. D., Uyen, P. T. H., & Van, P. T. T. (2021). Common Errors in Writing of EFL Sophomores in a Context of the Mekong Delta. *Eurasian Journal of Applied Linguistics*, 7(2), 46–57.
- Khatteer, S. (2019). An analysis of the most common essay writing errors among EFL Saudi female learners (Majmaah University). *Arab World English Journal*, 10(3), 364–381. <https://doi.org/10.24093/awej/vol10no3.26>.
- Ma'mun, N. (2016). The grammatical errors on the paragraph writing. *Jurnal Vision*, 5(1). <https://doi.org/10.21580/vjv5i1862>.
- Mai, N. T., & Ly, T. T. (2024). Current English education in Vietnam: Policy, practices, and challenges. In Nghia, T. L. H., Ly, T. T., & Mai, N. T. (Eds.), *English language education for graduate employability in Vietnam* (pp. 45–65). Springer. https://doi.org/10.1007/978-981-99-4338-8_3.
- Nadya, M., & Muthalib, K. A. (2021). Error analysis of the students' English written descriptive text. *English Education Journal*. <http://doi.org/10.24815/eej.v12i2.19552>.
- Nhu, P. Q., & Ho, P. V. P. (2024). Common syntactic errors in the writing skills of the 3rd year students at Van Lang University. *International Journal of TESOL & Education*, 4(2), 118–144. <https://doi.org/10.54855/ijte.24427>.
- Pradnyaswari, L. F. S., Udayana, I. N., & Mulyana, N. (2022). Syntactic errors made by English department students in writing undergraduate thesis abstract. *ALFABETA: Jurnal Bahasa, Sastra, dan Pembelajarannya*, 5(2), 57–65.
- Richards, J. C. (1970). A non-contrastive approach to error analysis. *English Language Teaching Journal*, 25(3), 204–219.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(1–4), 209–232. <https://doi.org/10.1515/iral.1972.10.1-4.209>.
- Setiaji, H., Fajar, M., & Sulistyowati, H. (2023). Syntactical errors of narrative essay made by EFL students low achievements. *JEELL (Journal of English Education, Literature and Linguistics)*, 9(2), 12–21.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732997>.

CÁC MẪU LỖ CÚ PHÁP TRONG BÀI VIẾT HỌC THUẬT CỦA SINH VIÊN NGÀNH NGÔN NGỮ ANH TẠI VIỆT NAM: DIỄN GIẢI DỰA TRÊN MỘT HỆ THỐNG PHÂN LOẠI KẾT HỢP

Võ Ngọc Ánh Minh¹

Diệp Vũ Đăng Quang² Nguyễn Trường Giang³

^{1,2}Trường Đại học Quốc tế, Đại học Quốc gia Thành phố Hồ Chí Minh; ³Trường Đại học Thành Đô

Email: vongocanhminhcontactwork@gmail.com¹; dangquangdiep0510@gmail.com²;

ntgiang@thanhdowni.edu.vn³

Ngày nhận bài: 01/11/2025; Ngày phản biện: 4/12/2025; Ngày tác giả sửa: 8/12/2025;

Ngày duyệt đăng: 25/12/2025

DOI: <https://doi.org/10.58902/tcnckhpt.v4i4.311>

Tóm tắt: Nhiều sinh viên ngôn ngữ Anh ở bậc đại học vẫn gặp khó khăn khi ứng dụng cú pháp trong viết học thuật, mặc dù đã được tiếp xúc đáng kể với các khóa học ngữ pháp chính quy. Nghiên cứu này xem xét hiện tượng mắc lỗi cú pháp trong nhóm sinh viên năm nhất chuyên ngành Ngôn ngữ Anh tại một trường đại học ở Việt Nam. Với phương pháp hỗn hợp, nghiên cứu phân tích 49 bài luận dạng “vấn đề – giải pháp” bằng cách sử dụng một hệ thống phân loại kết hợp, tích hợp độ chính xác trong phân loại của mô hình Surface Strategy Taxonomy (Dulay et al., 1982) với các phân biệt sự phạm của Ferris (2005). Kết quả cho thấy lỗi định dạng sai (misformation) chiếm ưu thế rõ rệt (48.36%), vượt xa lỗi lược bỏ (omission) (19.07%) và các loại lỗi khác. Diễn giải cùng dữ liệu khảo sát, nghiên cứu phát hiện lịch sử học tập nặng về giảng dạy theo hướng quy tắc của đối tượng tham gia nghiên cứu. Điều này gợi ý rằng những lỗi cú pháp quan sát được không xuất phát từ việc thiếu kiến thức ngữ pháp, mà xuất phát từ yêu cầu vận dụng ổn định kiến thức ngữ pháp trong các nhiệm vụ viết dài và có tính học thuật. Lỗi thiếu mạo từ - thường được nhắc tới trong các mô tả về ảnh hưởng của tiếng mẹ đẻ - xuất hiện đồng thời với những nỗ lực sử dụng cấu trúc phức tạp, cho thấy sự tương tác qua lại giữa nền tảng ngôn ngữ, trải nghiệm học tập và điều kiện của nhiệm vụ viết. Những phát hiện này nhấn mạnh giá trị của phân tích lỗi theo hướng chẩn đoán và gắn với bối cảnh, thay vì chỉ dừng lại ở thống kê mô tả, đồng thời gợi mở các định hướng sự phạm liên quan đến thực hành viết có ý nghĩa và sự chú ý phản tư trong mối quan hệ giữa hình thức và chức năng ngôn ngữ.

Từ khóa: Dạy học ngữ pháp theo quy tắc; Lỗi cú pháp; Lỗi định dạng sai; Liên ngôn ngữ; Phân tích lỗi theo hướng chẩn đoán; Viết học thuật.