Original Article

Lexical Errors in First-Year Undergraduates' English Compositions: An Error Analysis Approach

Mary Awino[1]* & Dr. Nancy Ikaria, PhD[1]

[1] Kenyatta University, P. O. Box 43844-00100, Nairobi, Kenya.

* Author for Correspondence ORCID ID; https://orcid.org/0009-0007-7202-983; Email: awinom312@gmail.com

ABSTRACT

This article explores the use of the Error Analysis (EA) approach in investigating lexical errors in the written English compositions of first-year undergraduate students. Lexical competence is a crucial component of academic writing proficiency, yet it remains a persistent challenge for second-language (L2) learners at the university level. The study employed a qualitative research design, which was descriptive in nature. Using a corpus of essays collected from 20 first-year undergraduates who were studying English Communication Skills for Professional Development, the study applied a structured EA framework to identify lexical deviations in the students' compositions. It then categorised the lexical deviations using the distinction between form-oriented and content-oriented lexical errors, which had nine subcategories. The study justifies the choice of EA over other analytical approaches by emphasising its learner-centred focus, practical applicability, and methodological clarity. While acknowledging critiques of EA—such as its limited focus on errors and potential subjectivity in error classification—the paper defends its relevance by drawing on the work of scholars who support its use in pedagogical research. Findings revealed that errors in word formation/morphology, as well as misspelling errors, were the most prevalent, highlighting the need for targeted vocabulary instruction. The study concludes by discussing the pedagogical implications of the findings and proposing directions for future research into lexical development and instructional design.

APA CITATION

Awino, M. & Ikaria, N. (2025). Lexical Errors in First-Year Undergraduates' English Compositions: An Error Analysis Approach. *East African Journal of Arts and Social Sciences*, *8*(3), 20-33. https://doi.org/10.37284/eajass.8.3.3391

CHICAGO CITATION

Awino, Mary and Nancy Ikaria. 2025. "Lexical Errors in First-Year Undergraduates' English Compositions: An Error Analysis Approach." *East African Journal of Arts and Social Sciences* 8 (3), 20-33. https://doi.org/10.37284/eajass.8.3.3391.

HARVARD CITATION

Awino, M. & Ikaria, N. (2025), "Lexical Errors in First-Year Undergraduates' English Compositions: An Error Analysis Approach", *East African Journal of Arts and Social Sciences*, 8(3), pp. 20-33. doi: 10.37284/eajass.8.3.3391.

INTRODUCTION

Mastery of vocabulary and word choice—often referred to as lexical competence—is a critical component of successful second language (L2) writing. For first-year undergraduates studying English Communication Skills for Professional Development – a course unit that focuses on improving students' writing skills – limited lexical proficiency can significantly affect written coherence, clarity, and overall communicative effectiveness. Among the various categories of linguistic errors observed in L2 writing, lexical errors are particularly disruptive as they can obscure meaning more than grammatical errors (Liu & Shaw, 2001).

This article examines the categories and frequency of lexical errors in the written compositions of first-year undergraduates and evaluates the effectiveness of the Error Analysis (EA) approach in identifying and categorising these errors. Lexical errors, which include confusion of two formally similar words; wrong word formation/morphology error; word invention; borrowing from the L1; adaptation of an L1 to the L2; linguistic calque; misspelling of certain linguistic components; incorrect selection of two semantically related words; and collocational errors (James, 1998) can offer valuable insights into learner development when analyzed systematically (Ellis, 1997).

Several methodologies have been employed to study learner language, including Contrastive Analysis (CA), Interlanguage Theory, and Corpus Linguistics. However, the EA approach was selected for this study because of its systematic framework, which focuses on actual learner output (Corder, 1967; Ellis, 1997; Granger, 1999). Unlike CA, which relies on comparing language systems, EA centres on the learner's internal language system as revealed through errors. This makes EA particularly suitable for examining lexical inaccuracies that emerge during the writing process.

This article aims to identify the most frequent types of lexical errors made by first-year undergraduates; evaluate the effectiveness of EA in uncovering patterns and causes of these errors; and reflect on the strengths and limitations of EA in light of both supportive and critical scholarly perspectives.

By applying EA to a corpus of student essays, the study contributes to the broader understanding of lexical development in L2 writing and provides pedagogical implications for vocabulary instruction at the tertiary level.

LITERATURE REVIEW

Lexical Errors in Second Language Writing

Lexical errors are among the most persistent and significant challenges for L2 learners, particularly in academic writing. According to Llach (2011), the most common lexical error categories found in literature are word choice errors, omissions, irregular word forms, word order, borrowings, lexical creations, and spelling. These errors often stem from a combination of limited vocabulary knowledge, L1 transfer, and overgeneralization (Zheng & Park, 2013). In academic writing contexts, such errors can affect not only comprehension but also perceptions of competence and credibility.

The Error Analysis Approach

Error Analysis, introduced by Corder (1967), provides a systematic way of examining the errors learners make as a window into their interlanguage development. The process typically involves four

stages: choosing the language corpus, identifying errors in the corpus, classifying errors, explaining errors, and evaluating errors. EA is especially useful in pedagogical contexts, where identifying recurring error types can inform teaching strategies and materials design (James, 1998b).

The strength of EA lies in its learner-centred focus and its ability to make inferences about internal language processes based on external performance. Unlike CA, which focuses primarily on differences between L1 and L2 (Lado, 1957), EA accounts for both interlingual and intralingual sources of error, thereby offering a more nuanced view of learner development.

Critiques of Error Analysis

Despite its utility, EA has been critiqued on several fronts. Selinker & Rutherford (2013) pointed out that EA often fails to capture the dynamic and variable nature of interlanguage development. One major concern is that EA emphasises what learners do wrong, while ignoring what they do right, potentially leading to a deficit-oriented view of learner language. Moreover, the classification of errors can be subjective, especially with lexical items, where meaning is often context-dependent (Spada & Lightbown, 2006).

Another critique concerns the assumption of native-like norms as the benchmark for correctness. This can obscure legitimate learner choices that deviate from the native norm but are still communicatively effective, a point raised in post-structuralist critiques of SLA research (Kramsch, 2012; Pennycook, 2004).

Defending the Use of EA

Despite these criticisms, many scholars argue for the continued relevance of EA, particularly in pedagogical research. James (1998a) emphasises that the practical benefits of EA in classroom settings outweigh its theoretical limitations. Gass *et al.* (2020) argue that when complemented with other approaches, EA can reveal important aspects of

learner strategy use and developmental stages. Furthermore, studies like Dulay (1982) demonstrate that EA is particularly effective in identifying frequent and fossilised error types that are resistant to correction.

In this study, EA was selected not only for its methodological clarity but also because of its proven effectiveness in analysing written learner output, especially in institutional academic settings. Its focus on actual learner production makes it especially appropriate for investigating lexical errors, which often require nuanced contextual interpretation.

METHODOLOGY

Research Design

This study employed a qualitative research design grounded in the principles of Error Analysis (EA) to examine lexical errors in the written compositions of first-year undergraduate students. EA, as outlined by S. Corder (1974) and James (1998a), involves a systematic process of identifying, classifying, and interpreting learner errors to reveal underlying linguistic or cognitive processes. The study was descriptive in nature and exploratory in scope, aiming to identify patterns of lexical errors and assess the efficacy of the EA approach in categorising these errors within a real-world academic context.

Participants

A hundred and seventy-seven (177) first-year undergraduate university students doing a Bachelor of Arts with Education were subject to composition writing. The researcher preferred these students because: (i) they were teachers in the making; therefore, it was appropriate to involve them in an LEA study to prepare them to be better teachers and or, communicators and (ii) they were doing a cross-cutting course unit called English Communication Skills for Professional Development which would therefore enable the researcher administer the instrument through a class test. A total of 20 scripts

were randomly sampled from 177 scripts by picking every 9th script until the 20 scripts were collected. Random sampling was used because each member of the larger population had an equal chance of selection, and it allowed the researcher to use a non-biased sample. This non-biasedness, in turn, led to truthful conclusions being drawn about the larger population (DePersio, 2018). It was evident that the sample was less than half of the target population. Nonetheless, the researcher's decision to choose a small sample size was justified by claims that handling and analysing big data sets might be challenging and repetitive. According to Labov (1966) and Trudgill *et al.* (1974), there is no longer a need for big samples, especially when it comes to collected data, as they do not guarantee the presence of noteworthy characteristics. Furthermore, as noted by Milroy (1987), language is generally more homogeneous than other human behavioural forms and requires less data for analysis because an excessive amount of it tends to show recurring patterns that exacerbate data management problems.

Data Collection and Analysis

The data comprised a set of 20 written essays, limited to not more than three pages. The essay was part of a regular in-class writing assignment under controlled and timed conditions. Students were asked to write an expository essay on 'the effects of COVID-19 on education in Uganda'. The essays were handwritten and anonymised for analysis.

In order to identify and sort out the students' lexical deviations, the researcher read the 20 compositions at least twice. In this study, lexical deviations were taken as a departure from the lexical standard. That is, they were considered incorrect forms or uses of vocabulary items on any basis, such as orthographic, morphological, syntactical, or semantic. The researcher then classified the identified lexical errors according to the distinction between form-oriented and content-oriented lexical errors which were nine in number, namely: confusion of two formally similar words; wrong

word formation/morphology error; word invention; borrowing from the L1; adaptation of an L1 to the L2; linguistic calque; misspelling of certain linguistic components through addition, omission, and misordering; incorrect selection of two semantically related words (near synonyms, and hyponyms instead of supernyms); and collocational errors.

Each of the 20 students' written compositions had a results sheet. Whenever a particular lexical error appeared in the student's composition, the researcher recorded it in the results sheet under the column 'identified error'. These results sheets were used to count the frequency of each lexical error. Repeated lexical errors were counted so that if a learner produced the same lexical error, it would be counted only once, e.g., a subject might write: *The lecture was kind to me. The lecture advised me to attend classes regularly. I will never forget that lecture for the rest of my life.* Here, only one lexical error, i.e., *lecture* is present. The total number of all the form-oriented and content-oriented categories of lexical errors was then ascertained to calculate percentages for each error committed. The frequency ratings were illustrated by means of tables showing frequency counts and percentages. The accurate identification of errors is a means to establish their causes (Carrió-Pastor & Mestre-Mestre, 2014). In other words, once the researcher identified the lexical errors, their identification pointed her to their causes.

Theoretical Framework

Error Analysis was used as the framework in this study. The study and analysis of second language learners' errors is known as Error Analysis (EA). S. Pit Corder and his colleagues established EA in the 1960s as a subfield of applied linguistics to demonstrate that many learner errors were caused by a variety of factors rather than the learner's mother tongue, grouped under the term intralingual errors.

Richards & Schmidt (2010) state that "an intralingual error results from faulty or partial learning of the target language, rather than from language transfer." The following are the types of intralingual errors:

- Overgeneralizations are errors caused by applying target language principles to inappropriate situations. For example, they can occur when dealing with regular and irregular verbs, as well as the use of plural forms (Richards *et al.*, 2002).

- Simplifications are linguistic rules that learners make either by omission or addition that are less complex than those found in the TL.

- Induced errors occur when the definitions, examples, explanations, and organisation of practice opportunities provided by the teachers are misleading to the students (James, 2013).

- Communication strategy is a method for a learner with low language skills to articulate a meaning in a second or foreign language (Richards & Schmidt, 2010). In trying to communicate, a learner may have to make up for a lack of knowledge of grammar or vocabulary through approximation – *near synonyms, superordinate terms, antonyms/opposites*; language switching from L2 to L1; use of calque; or by coining words (James, 1998b, 2013).

- Errors of avoidance occur when people fail to use specific target language structures because they are perceived to be too difficult. Therefore, they may use their first language (L1) as a substitute or find an alternative or approximative means of communicating their idea in the L2 (James, 1998).

- Overproduction errors come about when beginners regularly repeat a structure because they have not yet acquired and accumulated sufficient linguistic information to enable them

to employ the target language's finite rules to construct endless structures.

The Error Analysis approach focuses on choosing the language corpus, identifying errors in the corpus, classifying errors, explaining errors, and evaluating errors as explained below:

A. Choosing the language corpus: This is accomplished by gathering a sample of the learner's language, which serves as a foundation for error analysts to use in determining their desired outcomes. "The data that are concerned, in this case, are typically written, although oral data can be used" (Gass, 2013). The data reveals how ESL learners employ TL in their writing or speech. In the current study, the data collected were from students' written English compositions.

B. Identifying the error in the corpus: The research analyst only looked for lexical errors that were specifically orthographic deviations, ignorance of syntactic restrictions, and meaning deviations.

C. Classifying error: "Error analysts have to classify the recognised errors in terms of how learners produce the TL differently from its native speakers." (Arb-Chaaba, 2017). The current study used the distinction between form-oriented and content-oriented lexical errors criterion as explained below:

The distinction between form-oriented and content-oriented lexical errors criterion has its justification in the observation that the mental lexicon is organised following both formal and semantic principles (Meara, 1996a). This criterion is widely used, captures a substantial number of lexical errors, and is used by many studies to classify lexical errors (Arb-Chaaba, 2017; Hemchua & Schmitt, 2006; James, 1998b; Mutta, 1999; Rezai & Davarpanah, 2019; Zimmermann, 1987). Lexical error types under this classification criterion are explained below according to James (1998):

1. Confusion of two formally related words: These are word pairs (or triples) that look and sound alike: *accessory/accessory, affect/effect, Mary/marry/merry, prize/price, ledge/pledge, save/safe, three/tree, fry/fly, flea/free, glow/grow, fresh/flesh, reserve/preserve etc).*

2. Wrong word formation/morphology error: these errors occur when a part of the five types of lexical words – noun, verb, adjective, adverb, and preposition – is not supplied according to the norm: *six book*, aboli*shment*($\sqrt{}$ -tion) for nouns; **bringed, was drink*en*($\sqrt{}$ -ing) for verbs; *visit me soon* ly* for adverbs; and *a colourfull*er scene,* for adjectives. Prepositions happen to have no morphology.

3. Word invention - Some errors result in 'terms' that don't exist in the FL/TL. They can either come from the MT or be generated by the learner using the TL's resources.

4. Borrowing from the mother tongue without making any changes to the word.

5. Coinage or adaptation of an L1 to the L2 or the new expression, which is derived from L1, is modified to the TL's structure.

6. Linguistic calque – if the L2 word is a direct translation of an L1 word.

7. Misspelling of certain linguistic components through: (a) Omission, for example, Mahmoud (2014) gives examples: **no (know), *dout (doubt), *weit (weight), *intresting (interesting),* etc. (b) Addition or over inclusion, e.g., *fresh(er)men, din(n)ing room, ti(t)tle,* and many others. (c) Misordering: e.g., *littel (little), ferporate* (perforate)

8. Incorrect selection of two semantically related words.

   ▪ A pair (or more) of words that are semantically related yet distinguishable are called near or assumed synonyms. Examples are: *malice/spite, coat/jacket,*
   chemist/apothecary/pharmacist, big/huge, snake/serpent, hen/cock, etc. She knew he was a thug, but she married him. What *mistake ($\sqrt{}$blunder)!

   ▪ Hyponyms instead of supernyms category, for example: using a general term when a more precise one is required, which results in a lack of sufficient elaboration of the concept: *She is *good ($\sqrt{}$exceptional/$\sqrt{}$reliable/$\sqrt{}$reputable).*

9. Collocational errors: Collocations are the other words with which a specific word is frequently associated (James, 2013). For example, we can say that an army has suffered *big losses,* but *heavy losses* are preferred. We *make an attempt* and *have a try,* but can neither **make a try* nor **have an attempt,* despite the synonymy of *attempt/try.* We also have irreversible binomials like *fish and chips* (not **chips and fish)* and *strawberries and cream.*

D. Explaining error: This is regarded as the most important step (Arb-Chaaba, 2017). It accounts for why the identified errors have been made, hence determining their sources. The form-oriented and content-oriented lexical errors were categorised as interlingual or intralingual based on their identification.

E. Evaluating error: The most commonly used criterion through which errors are to be judged is gravity or seriousness (Ellis & Barkhuizen, 2005). In the current study, the errors that impeded meaning the most were branded very serious. A list of the errors identified in the students' compositions was organised in a hierarchy of – most serious, serious, and least serious – depending on how much they impeded meaning (Endley, 2016).

The severity of an error is usually determined by the reader's attitude and cannot be determined by a universal standard (Nushi *et al.,* 2021). Therefore, error gravity was determined by the researcher since

she was the one analysing the 20 compositions. The more lexical errors found in a student's work, the worse his or her writing.

Therefore, EA was used in the proposed study because it is a reliable approach.

The third phase of the Error Analysis approach is to classify the errors that have been identified in phase two. Data was categorised using one of Llach's (2011) taxonomies, the distinction between form-oriented and content-oriented lexical errors criterion. This criterion is very frequent, captures a large number of lexical errors, and is used in a variety of research to categorise lexical errors. According to Meara (1996), this criterion asserts that the mental lexicon is organised following both formal and semantic principles. James (1998) breaks down the distinction between form-oriented and content-oriented lexical errors criterion into nine categories, namely: [1]confusion of two formally similar words; [2]wrong word formation/morphology error; [3]word invention; [4]borrowing from the L1; [5]adaptation of an L1 to the L2; [6]linguistic calque; [7]misspelling of certain linguistic components; [8]incorrect selection of two semantically related words; and [9]collocational errors.

Limitations

While the EA framework provided a structured and replicable method for analysis, some limitations were acknowledged:

- The classification of lexical errors often involves subjective judgment.

- The scope was limited to a single institution and one writing task type.

- The analysis did not include spoken data, limiting generalizability across language modalities.

RESULTS

This section presents the findings from the analysis of 20 written compositions by first-year undergraduate students. A total of 637 lexical errors were identified across all the essays, and these were categorised according to the error types defined in the framework. The frequency and percentage distribution of each error type are presented in Table 1.

Frequency and Types of Lexical Errors

Table 1: Distribution of Lexical Error Types

S/No	Error Category	Frequency	Percentage	Example of Error	Correction
1.	Wrong word formation/morphology error	335	52	Forexample	For example
2.	Misspelling of certain linguistic components	161	25	droped	dropped
3.	Incorrect selection of two semantically related words	81	13	… many girls *acquire pregnancies …	… many girls **get** pregnancies …
4.	Confusion of two formally related words	26	4	ware	were
5.	Collocational errors	16	3	dependant source	source of income
6.	Linguistic calque	13	2	… and they were later manufacturing other medicine to *chase away Corona virus …	… and they were later manufacturing other medicine to **eliminate/prevent / cure)** Corona virus …
7.	Word invention	05	1	tangets	detergents

S/No	Error Category	Frequency	Percentage	Example of Error	Correction
8.	Adaptation of an L1 to the L2	00	00	00	00
9.	Borrowing from the L1	00	00	00	00
Total Frequency & Percentage		**637**	**100%**		

Most and Least Common Errors

The study found that the wrong word formation/morphology error scored the highest frequency of 335 (52%). These results indicate that the majority of students have difficulty in forming words. The morphology error was broken down into nine sub-categories numbered a-i because students were forming words erroneously in different ways or using the wrong form of the word; that is, they were: [a] separating two lexical words that should be written as one; [b] separating root words from their affixes or combining them with the wrong affix; [c] writing two words as one, yet they should be written separately; [d] writing verbs in the wrong tense or those which are inappropriate not necessarily because of tense; [e] adding/omitting plural morphemes, or using a singular instead of the plural form of the word and vice versa; [f] adding or omitting morphemes which changed the grammatical function or meaning of the word; [g] using the wrong article; [h] using the wrong preposition; and lastly, [i] using American instead of British vocabulary.

Misspelling of certain linguistic components, with a frequency of 161 (25%), came second to morphology errors. They were further broken into five categories, namely: [1]omission; [2]addition or over-inclusion; [3]misordering; [4]omission and addition of lexical components in the same word; and [5]omission and misordering of lexical components in the same word.

The incorrect selection of two semantically related words with a frequency of 81 (13%) took third place. It was broken down into two sub-categories: near or assumed synonyms and hyponyms instead of supernyms. The following error categories received a low frequency: confusion of two formally related words occurred 26 times (4%); collocational errors occurred 16 times (3%); linguistic calque occurred 13 times (2%); and word invention occurred 05 times (1%).

The error categories that received zero frequency in all 20 compositions include adaptation of an L1 to the L2 and borrowing. The morphology error topped the list of the most serious error category because it received the highest frequency, while the least serious categories were adaptation of an L1 to the L2 and borrowing from the L1 because they received zero frequency in all 20 WECs. WEC-H got the highest frequency of LEs (58) and therefore it was branded the worst written, while WEC-N received the lowest frequency (13), making it the best written.

DISCUSSION

The analysis of lexical errors in the written compositions of first-year undergraduates revealed several key patterns that align with and extend previous research in second language acquisition (SLA) and error analysis. The two most prominent error types – morphology error and misspellings – together accounted for 70% of all the errors, suggesting that these areas remain major stumbling blocks for L2 learners even at university level.

Interpretation of Key Error Types

The morphology or wrong word formation errors, which received the highest frequency, were broken down into the following nine parts:

- Words that combine two lexical words to form compounds were separated (hence altering their meaning), yet they should be written as one, for example, the word *a lot* should be written with that space between the article a

- Words that were separated from their affixes or combined with the wrong affix

- Words that were written as one, yet they should be written separately

- Verbs that were written in the wrong tense or used inappropriately

- Errors that were made by adding/omitting plural morphemes, or using a singular instead of the plural form of the word, and vice versa

- Words that were formed by adding or omitting morphemes that changed the grammatical function or meaning of the word

- Wrong article

- Wrong preposition choice

- Use of American instead of British vocabulary

Misspellings, which were the second most frequent, further highlight a lack of familiarity with a variety of vocabulary in English and also careless writing. Some students do not bother to proofread their write-ups. They just hand in rough work. Misspellings are often not addressed directly in language instruction. Their prevalence supports previous findings by researchers such as Nyamasyo (1994), who noted that errors of spelling rank first amongst all the different types of grammatical and lexical errors recognised from the language performance data from native English college-level students and from those for whom English is a second language.

The occurrence of synonym misuse under the category – *incorrect selection of two semantically related words* – came in third place with 13% frequency. This occurrence indicates that learners may be applying word formation rules or depending on presumptive equivalencies without fully grasping context or connotation. These kinds of mistakes are frequently developmental and can be a sign of a desire to take lexical risks, which is good

for learner initiative but problematic if not accompanied by sufficient vocabulary.

Sources of Lexical Errors

Consistent with the principles of Error Analysis, the lexical errors observed in this study can be attributed to the following main sources:

95% of the LEs in the students' WECs were intralingual, while only 5% were interlingual. Intralingual errors occur as a result of faulty or incomplete target language learning rather than language transfer. Out of the six types of intralingual errors, three were seen as the underlying cause of the LEs in the students' WECs. They include: overgeneralizations, simplifications, and communication strategies.

On the other hand, the errors caused by mother tongue influence were not many. They received 5% which is a clear indication that most of the LE students' mastery is not a result of their mother tongues. Therefore, the L1 should not be blamed entirely for the LEs students make in their WECs. The L1 does more good than harm to a learner's vocabulary acquisition. No wonder Uganda shifted to a thematic curriculum approach, which encourages pupils to learn in their native languages for the first three years of primary school before switching to English in the fourth (Lamwaka, 2020).

These findings support the utility of EA in uncovering the cognitive and linguistic processes behind learner output, especially when the focus is on written language, where errors can be documented and analysed.

Evaluating the Use of Error Analysis

The application of the EA approach in this study proved to be both productive and pedagogically relevant. Its structured methodology enabled the systematic identification and categorisation of lexical errors, while its emphasis on learner-produced data ensured a grounded, authentic view of language use in a real-world academic context.

However, several critiques of EA must also be acknowledged and addressed:

Schachter & Celce-Murcia (1977) argued that traditional error analysis (EA) focuses only on overt learner errors (for example, grammar mistakes), ignoring important non-errors like avoidance (when learners avoid structures they can't use correctly). A learner might produce grammatically correct but overly simple sentences because they avoid complex syntax. EA would wrongly assume the learner has no problem, since there's no error on the surface. They add that EA fails to account for the communicative intent behind utterances, thus missing deeper issues in language development. They conclusively assert that EA gives a distorted view of learner interlanguage by neglecting what's not said, and therefore is inadequate as a complete model of L2 development.

Like Schachter & Celce-Murcia, Richards (1980) noted that EA tends to focus on linguistic form (syntax, morphology) at the sentence level, ignoring discourse-level errors or pragmatic competence. EA describes errors *after* they occur but doesn't predict what kinds of errors learners might make, making it reactive, not explanatory, or developmental. The approach underplays the idea that learner language is a systematic, evolving interlanguage, not just a set of mistakes. He concludes by saying that while EA is useful for diagnosing problems, it's not enough for understanding why errors occur or how they relate to broader L2 development.

Ferris (2006) focused on written corrective feedback, questioning whether marking only surface-level errors (like spelling or verb tense) leads to lasting improvement. She found that while students may fix surface-level mistakes short-term, the development of syntactic complexity and fluency often lags. She added that teachers should distinguish between treatable (surface) and untreatable (deeper structural) errors and not overwhelm learners by correcting everything. Effective error correction requires a strategic

approach—EA alone is insufficient for guiding pedagogical decisions.

Gass & Selinker (1992) noted that EA often emphasises errors at the expense of correct language use, which can create a deficit-oriented perception of learner competence. In this study, this concern was mitigated by analysing lexical errors in context, recognising instances of partial or near-correct usage, and acknowledging learners' communicative intentions.

Another particular concern in the analysis of lexical errors is subjectivity in error categorisation, where semantic boundaries are often fuzzy. To address this, inter-rater coding was used to enhance reliability, and ambiguous cases were discussed collaboratively (Larsson *et al.,* 2020).

Additionally, EA often presumes native-like norms as the standard for correctness, which can marginalise acceptable non-standard usage. While this study used academic English as the referent (given the institutional context), the analysis was sensitive to alternative but intelligible lexical choices that still served the communicative purpose.

Despite these limitations, scholars such as James (1998b) and Selinker & Gass (2008) argue that EA retains its value, particularly in contexts where the goal is instructional improvement. This study supports that view by demonstrating how EA can reveal patterns of lexical error that inform more targeted teaching strategies, particularly in vocabulary instruction and feedback mechanisms.

Pedagogical Implications

The findings suggest several practical steps for improving lexical competence among undergraduate learners as explained below:

Explicit instruction, especially on morphology, spelling, and synonyms, should be incorporated into academic writing courses. Lexical awareness activities that differentiate near-synonyms and address common word choice confusions can

enhance precision in writing. Data-driven learning, including learner corpora and concordance tools, may help students become more self-aware of word use in context. Instructors should also provide more context-sensitive feedback that goes beyond marking lexical errors as simply "wrong," instead explaining why a particular choice may be inappropriate or awkward, and offering more natural alternatives.

Summary

Overall, the use of Error Analysis in this study provided a detailed and nuanced picture of lexical errors among first-year undergraduates. While the approach is not without its limitations, its ability to systematically identify and interpret patterns of learner error makes it a valuable tool for both researchers and educators. By critically applying EA and engaging with both its strengths and critiques, this study contributes to ongoing conversations about L2 lexical development and pedagogically responsive language instruction.

CONCLUSION

This study set out to examine the effectiveness of the Error Analysis (EA) approach in identifying and categorising lexical errors in the written English compositions of first-year undergraduate students. Through a detailed analysis of 20 written English compositions, the study found that morphology/word formation errors were the most frequent types of lexical misuse, followed by misspellings, incorrect selection of two semantically related words, and confusion of two formally related words. These patterns point to ongoing challenges in L2 learners' lexical development, especially in contexts that demand precision and fluency in academic writing.

The use of EA as the primary analytical tool was both justified and effective. Its structured methodology enabled a systematic identification of errors and facilitated insights into underlying learner processes, such as L1 interference and developmental overgeneralization.

While concerns surrounding EA, such as its focus on errors and subjectivity in classification, are valid, this study addressed them by categorising the data using one of Llach's (2011) taxonomies, the distinction between form-oriented and content-oriented lexical errors criterion, which has its justification in the observation that the mental lexicon is organised following both formal and semantic principles (Meara, 1996a). The advantages of using this criterion are that it is widely used, captures a substantial number of lexical errors, and is used by several studies to classify lexical errors (Arb-Chaaba, 2017; Hemchua & Schmitt, 2006; James, 1998b; Mutta, 1999; Rezai & Davarpanah, 2019; Zimmermann, 1987). The nine lexical error types under this classification criterion were explained according to James (1998). In doing so, it demonstrated that EA, when applied critically, remains a valuable tool in both research and pedagogical contexts.

The findings hold several pedagogical implications. Vocabulary instruction at the tertiary level should prioritise not only breadth of vocabulary but also depth, focusing on word formation, spelling, lexical appropriateness, and register. Instructors should also consider integrating corpus-based tools, contextualised feedback, and lexical awareness activities to help learners internalise word usage patterns more effectively.

This research contributes to the broader discourse on lexical competence in second language writing and reaffirms the relevance of EA in understanding learner language. However, it is not without limitations. The scope was restricted to a single writing task within one institution, and the findings may not generalise across other learner populations or genres. Future research could adopt a longitudinal design to track lexical development over time or compare EA with other analytical frameworks, such as learner corpus analysis or interlanguage pragmatics.

In conclusion, EA remains a powerful approach for diagnosing learner difficulties, especially in lexical

use. When applied thoughtfully, it offers rich insights that can inform curriculum development and enhance the effectiveness of academic writing instruction.

REFERENCES

Arb-Chaaba, R. (2017). *Investigating Learners' Lexical Errors in Writing Skill* [Working Paper]. http://dspace.univ- guelma.dz:8080/xm lui/handle/123456789/2414

Carrió-Pastor, M. L., & Mestre-Mestre, E. M. (2014). Lexical errors in second language scientific writing: Some conceptual implications. *International Journal of English Studies*, *14*(1), 97–108.

Corder, S. (1974). *Error Analysis and Remedial Teaching.* https://eric.ed.gov/?id=ED101573

Corder, S. P. (1967). The Significance of Learner's Errors. *IRAL - International Review of Applied Linguistics in Language Teaching*, *5*(1–4). https://doi.org/10.1515/iral.1967.5.1-4.161

DePersio, G. (2018, February 13). *The Advantages of Using Simple Random Sample to Study Larger Populations*. Investopedia. https://www.investopedia.com/ask/answers/04 2915/what-are-advantages-using-simple-random-sample-study-larger-population.asp

Dulay, H. (1982). *Language two.* ERIC. https://eric.ed.gov/?id=ED227699

Ellis, R. (1997). *SLA Research and Language Teaching.*

Ellis, R., & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford University Press.

Endley, M. (2016, January 1). Perceptions of Error Gravity among Native Speaking and Non-native Speaking Faculty at UAEU. *Proceedings of the 21st TESOL Arabia Conference*. Proceedings of the 21st TESOL Arabia Conference, Dubai.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short-and long-term effects of written error correction. *Feedback in Second Language Writing: Contexts and Issues*, *81104*. https://www.cambridge.org/core/services/aop-cambridge- core/content/view/78F00AFBCC4 CFCD4ACD6B19783B4F2A6

Gass, S. M. (2013). *Second language acquisition: An introductory course*. Routledge.

Gass, S. M., Behney, J., & Plonsky, L. (2020). *Second language acquisition: An introductory course*. Routledge. https://www.taylorfrancis.c om/books/mono/10.4324/9781315181752/seco nd-language-acquisition-susan-gass-jennifer-behney-luke-plonsky

Gass, S. M., & Selinker, L. (1992). *Language transfer in language learning: Revised edition* (Vol. 5). John Benjamins Publishing. https://books.google.com/books?hl=en&lr=&i d=u02cvCT9krAC&oi=fnd&pg=PP2&dq=Seli nker+(1992)+&ots=5- WWhsX2tX&sig=w_T R6XTiHeWMG2h9XHkF582vOi0

Granger, S. (1999). Uses of tenses by advanced EFL learners: Evidence from an error-tagged computer corpus. *Language and Computers*, *26*, 191–202.

Hemchua, S., & Schmitt, N. (2006). *An analysis of lexical errors in the English compositions of Thai learners*.

James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*.

James, C. (1998b). *Errors in Language Learning and Use: Exploring Error Analysis*. Longman.

James, C. (2013). *Errors in language learning and use: Exploring error analysis*. Routledge.

Kramsch, C. (2012). Authenticity and legitimacy in multilingual SLA. *Critical Multilingualism Studies*, *1*(1), 107–128.

Labov, W. (1966). *The Social Stratification of English in New York City. Washington, DC: Center for Applied Linguistics, 1966. 2006.: Cambridge/Cambridge U. Press. The Study of Nonstandard English. Washington, DC: National Council of Teachers of English, 1969. The Study of Non-Standard English. Champaign, IL: National Council of Teachers of*.

Lado, R. (1957). *Linguistics across cultures; applied linguistics for language teachers*. http://thuvienso.thanglong.edu.vn/handle/TLU/6837

Lamwaka, B. (2020, March 15). *Uganda Shifts Toward Teaching Students in Local Languages*. Global Press Journal. https://globalpressjournal.com/africa/uganda/uganda-shifts-toward-teaching-students-local-languages/

Larsson, T., Paquot, M., & Plonsky, L. (2020). Inter-rater reliability in Learner Corpus Research: Insights from a collaborative study on adverb placement. *International Journal of Learner Corpus Research*, *6*(2), 237–251. https://doi.org/10.1075/ijlcr.20001.lar

Liu, E. T. K., & Shaw, P. M. (2001). *Investigating learner vocabulary: A possible approach to looking at EFL/ESL learners qualitative knowledge of the word*. *39*(3), 171–194. https://doi.org/10.1515/iral.2001.001

Llach, M. del P. A. (2011). *Lexical errors and accuracy in foreign language writing*. Multilingual Matters.

Meara, P. (1996). The dimensions of lexical competence. *Performance and Competence in Second Language Acquisition*, *35*, 33–55.

Meara, P. (1996b). The dimensions of lexical competence. *Performance and Competence in Second Language Acquisition*, *35*, 33–55.

Milroy, L. (1987). *Observing and analysing natural language: A critical account of sociolinguistic method*. Basil Blackwell.

Mutta, M. (1999). The role of vocabulary in language learning: A study of entrance examination compositions. *Diálogos Hispánicos*, *23*, 333–348.

Nushi, M., Jafari, R., & Tayyebi, M. (2021). Iranian Advanced EFL Learners' Perceptions of the Gravity of Their Peer Written Lexical Errors: The Case of Intelligibility and Acceptability. *English Language Teaching*, *01*(01), 17.

Nyamasyo, E. A. (1994). An analysis of the spelling errors in the written English of Kenyan pre-university students. *Language, Culture and Curriculum*, *7*(1), Article 1. https://doi.org/10.1080/07908319409525167

Pennycook, A. (2004). Critical applied linguistics. *The Handbook of Applied Linguistics*, 784–807.

Rezai, M. J., & Davarpanah, F. (2019). The Study of Formal and Semantic Errors of Lexis by Persian EFL Learners. *International Journal of Cognitive and Language Sciences*, *13*(1), 41–47.

Richards, J. C. (1980). Second language acquisition: Error analysis. *Annual Review of Applied Linguistics*, *1*, 91–107.

Richards, J. C., Schmidt, R., Platt, H., & Schmidt, M. (2002). *Longman Dictionary of Language Teaching & Applied Linguistics* (3rd edition edition). Longman.

Richards, J. C., & Schmidt, R. W. (2010). *Longman Dictionary of Language Teaching and Applied Linguistics*. Routledge.

Schachter, J., & Celce-Murcia, M. (1977). Some reservations concerning error analysis. *Tesol Quarterly*, 441–451.

Selinker, L., & Gass, S. M. (2008). Second language acquisition. *Lawrence Erlhaum Ass*.

https://www.researchgate.net/profile/Larry-Selinker/publication/229078359_Second_Language_Acquisition_An_Introductory_Course/links/5592ae3c08ae1e9cb4296f95/Second-Language-Acquisition-An-Introductory-Course.pdf

Selinker, L., & Rutherford, W. E. (2013). *Rediscovering interlanguage*. Routledge. https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.4324/9781315845685&type=googlepdf

Spada, N., & Lightbown, P. M. (2006). *How languages are learned*. Oxford University Press.

Trudgill, P. (1974). *The social differentiation of English in Norwich* (Vol. 13). CUP archive.

Zheng, C., & Park, T.-J. (2013). An Analysis of Errors in English Writing Made by Chinese and Korean University Students. *Theory & Practice in Language Studies (TPLS)*, *3*(8). https://www.researchgate.net/profile/Sofe-Ahmed/publication/274661437_The_Current_Practices_of_Teaching_Grammar_in_CLT_at_Secondary_School_Level_in_Bangladesh_Problems_and_Probable_Solutions/links/5d0252ed4585157d15a712b5/The-Current-Practices-of-Teaching-Grammar-in-CLT-at-Secondary-School-Level-in-Bangladesh-Problems-and-Probable-Solutions.pdf#page=46

Zimmermann, R. (1987). Form-Oriented and Content-Oriented Lexical Errors in L2 Learners. *IRAL*, *25*(1), 55–67.