European Journal of English Language Teaching
ISSN: 2501-7136
ISSN-L: 2501-7136
Available on-line at: www.oapub.org/edu

DOI: 10.46827/ejel.v11i4.6913 Volume 11 | Issue 4 | 2026

NOTICING INTERLINGUAL LEXICAL ERRORS IN TURKISH EFL WRITING: A COMPARISON OF LEARNER, TEACHER, AND CHATGPT ERROR FLAGGING AND FEEDBACK TYPES

Funda Dündar[i]
Department of Informatics,
Hatay Mustafa Kemal University,
Türkiye
orcid.org/0000-0002-7209-5198

Abstract:
Written corrective feedback (WCF) from teachers and, more recently, from artificial intelligence (AI) tools has been widely studied; however, little attention has been paid to comparing how these sources address interlingual lexical errors. This study primarily aims to examine how the learners, the teacher, and AI feedback generated by ChatGPT differ in noticing and addressing interlingual lexical errors in EFL university writing. The present study was carried out at a state university where the English preparatory program introduces systematic writing instruction, progressing from paragraph writing (B1) to essay writing (early B2). Adopting a descriptive case study design, the study analysed 18 opinion texts produced by 12 B1-level students enrolled in an English preparatory programme at a Turkish state university. Six students each produced one opinion paragraph, while six different students each produced two opinion essays. A total of 81 error instances, identified through retrospective learner self-flagging, teacher feedback, and AI feedback, were classified according to James's (1998) taxonomy, and the feedback types were coded following Ellis's (2009) typology. The results showed that semantic errors constituted the majority of the instances (80.2%). The AI tool noticed the largest share of the errors (85.2%), followed by the learners (45.7%) and the teacher (28.4%), while all three sources converged on only 11 instances (13.6%). The teacher mostly provided indirect feedback with error codes, whereas the AI tool almost always gave direct feedback with an explanation. In this sense, it can be said that error identification works best as a complementary process, in which learner awareness, teacher judgement, and AI assistance each cover a different part of the errors.

Keywords: interlingual errors; lexical errors; written corrective feedback; ChatGPT; EFL writing

[i] Correspondence: email fundadundar@eskisehir.edu.tr

1. Introduction

Writing in a foreign language is linguistically demanding. It requires command of both grammar and vocabulary, which makes linguistic accuracy central to EFL writing. Writing in a foreign language requires a traceable and enduring record of learners' linguistic development, making it one of the most measurable skills through which progress in grammar, vocabulary, and accuracy can be directly observed (Williams, 2012; Son, 2024).

In some CEFR-aligned English preparatory programs EFL learners are expected to produce simple connected texts on familiar topics (Council of Europe, n.d., p. 5). Consistent with this CEFR specification, the English preparatory program where the present study was conducted introduces systematic paragraph-writing instruction from paragraph writing at the B1 level and essay writing at the beginning of the B2 level. At this stage of writing development, however, learners may experience difficulties in producing accurate English sentences. It is observed that many learners rely heavily on their mother tongue, Turkish (L1) system, when constructing sentences in English, which results in syntactic and lexical errors rooted in negative transfer (Hasırcıoğlu & Öztürk, 2024). In this context, written corrective feedback (WCF) is regarded as a key pedagogical tool for raising learners' awareness of such errors and encouraging the replacement of L1-induced forms with accurate target-language use (Makino, 1993). However, practical problems such as time limitation during feedback sessions, restricted space for writing drafts, and frequent misalignment between what teachers intend and what they actually do result in insufficient teacher feedback (Mao & Crosthwaite, 2019). In addition to these delivery-side problems, it was found that students, particularly those at particular lower proficiency levels, tend to lose focus on teacher-written corrective feedback and often fail to understand or apply the corrections they receive (Ferris *et al.*, 2011; Han & Hyland, 2015; Zheng *et al.*, 2018). Learners receiving teacher feedback may still need help to answer their questions about the ambiguity they experienced. In other words, learners may struggle to interpret the teacher feedback (Arts *et al.*, 2016). In this sense, self-correction sessions appear to be a necessary complement to teacher feedback, as they encourage learners to actively notice and understand their own errors in a more autonomous manner (Makino, 1993; Pham, 2022).

Recent advances in artificial intelligence (AI) technology have opened new possibilities for supporting learners in improving writing skills. Compared to teacher feedback, which may focus exclusively on accuracy, AI-based feedback assists learners with more detailed feedback, including content and organization-related parts of their texts (Al-Olimat, 2015). Because computerized feedback is convenient, learners may ask for more detailed feedback to improve accuracy and understand errors or enrich content anytime and anywhere they need (Liu, 2023). While comparative studies of teacher and AI-generated feedback have begun to emerge in EFL contexts (Wang, 2024), recent scoping reviews indicate that the field's findings as to the efficiency of feedback type (AI and/or Teacher feedback) remain incomplete and that significant gaps persist

(Crosthwaite & Sun, 2025). Interlingual errors have been documented as a persistent challenge for Turkish EFL learners (Gök, 2020; Hasırcıoğlu & Öztürk, 2024). In addition, a growing body of research has compared teacher and AI-generated feedback on EFL writing (Kaliisa *et al.*, 2026; Guo & Wang, 2024), and some studies have explored feedback practices in Turkish university preparatory schools (Han *et al.*, 2022; Er *et al.*, 2024). However, we have not encountered any study comparing teacher and AI-generated feedback on EFL writing that specifically focuses on interlingual lexical errors. This study primarily aims to examine how the learners, the teacher and AI-generated feedback differ in noticing and addressing interlingual lexical errors in EFL university writing. Exploring how learners, teachers, and AI-generated feedback differ in noticing such errors will help language teachers address students' linguistic needs and develop more productive writing feedback practices.

Within this framework, the research question is specified as follows:

1) To what extent do the learners, the teacher, and the AI tool differ in noticing interlingual lexical errors in EFL preparatory students' writing, and what types of feedback do the teacher and the AI tool provide for these errors?

2. Literature Review

Examining language learners' errors in second-language research has been a dynamic area of linguistic inquiry at various points in time. Corder (1967) describes learners' language as an idiosyncratic dialect, noting that both the target and second languages share a language system, and characterizes it as personal, unstable, and developmental. He also adds that errors are *the result of the first language* while learning a second language. James (1998) defines errors as unsuccessful bits of language. Ferris (2011) focuses on deviations from the rules of the target language. These definitions are based on two key characteristics: learner language is both developmental and influenced by systematic deviations from target-language norms.

Researchers, therefore, turned their attention to identifying the conditions under which errors arise, as understanding the origins of errors provides valuable insight into both theorizing second language development and designing pedagogical interventions. Richards (1971) describes the origins of errors as *intralingual*, which emerge from the learner's imperfect construction of the target language system. Accordingly, intralingual errors are frequently made but do not emerge from the background of the learners' mother tongue. In other words, they reflect the learners' competence at a specific stage. The other type of source is *interlingual*, which relates to the "*language produced by a nonnative speaker of a language*" (Gass & Selinker, 2001). Language learners depend on their mother language system to make sense of the target language; therefore, most errors result from a word-by-word translation strategy (Bhela, 1999). Nonnative speakers rely on structures assumed to be equivalent in their mother tongue, a phenomenon known as *interlanguage transfer*. James (1998) indicates that "*the clearest proof of L1 interference is where L1 nonstandard dialect features get transferred to L2*."

Classifying errors on syntactic and lexical levels is a common approach in Second Language Research. Syntactic errors typically stem from overgeneralization of language use or omission of a structure in the language system. Lexical errors, on the other hand, involve the inappropriate use of vocabulary, including the misuse of collocational patterns. A further well-established distinction is made between intralingual and interlingual error categories. Intralingual errors result from learners' inaccurate use of target language structures independent of mother tongue influence. They reflect the learner's language system by providing insightful evidence about the language competence or proficiency level of learners. As for interlingual errors, they arise from transfer between the mother tongue and the target language. The extent of transfer varies among individuals since learning strategies and processes are changeable. Comprehension of learner language is largely based on appropriate vocabulary use. Syntactical errors do not affect meaning as much as lexical errors. Native speakers identify non-native speakers from inappropriate word selection. For example, collocational knowledge is difficult to acquire and is a parameter of nativeness (Schmitt, 2000, as cited in Dodigovic *et al.*, 2017). Large corpora studies have shown that lexical errors constitute many learners' errors (Gass & Selinker, 2001). From this perspective, giving and receiving feedback on lexical errors should be prioritized, since lexical items carry meaning, and the inappropriate use of vocabulary leads to breakdowns in communication. From this perspective, the present study narrows its scope specifically to lexical errors.

2.1. Lexical Error Analysis

2.1.1 Formal Errors of Lexis

Within James's (1998) taxonomy, formal errors of lexis are those in which the deviation lies in the form of the word rather than its meaning. Two main sub-types are identified: formal misselection and misformation.

Formal misselection refers to errors of the malapropism type, or words likely to be confused due to orthographic and phonological similarity. These words are similar to the intended target word in terms of syllable count or initial sounds; however, they do not necessarily need to have similar meanings. To illustrate, in the sentence 'The Sahara dessert is huge,' the word *dessert* is mistakenly used instead of *desert*, despite the two words being phonologically and orthographically similar but semantically unrelated.

Misformation error types are related to non-existent words in the target language. These errors occur in three ways. Firstly, learners borrow words directly from their mother tongue and transfer them into the target language without adaptation. For instance, the Turkish word pasta means birthday cake, as in the sentence "*She blew the candles on the pasta*." The second type is coinage. Learners derive a new word from L1 and adapt it to the target language using word-formation patterns. For example, sportif is formed from Turkish, but the correct equivalent is sportive. The final misformation type is calque, which results from direct translation from L1 into the target language. To

exemplify, *bebek bakıcısı* is a Turkish phrase that is directly translated as *baby carer*, but in the target language, *babysitter* is used.

James (1998) also includes distortions, which he defines as "*the intralingual errors of form created without recourse to L1 resources*" (p. 150). Since the scope of the present study is limited to interlingual errors, distortions fell outside its framework.

2.1.2 Semantic Errors of Lexis

James (1998) describes semantic errors in lexis as errors in which learners use words that do not fit in the target language in context and intended meaning. Confusion of sense relations and collocational errors are two main types of semantic errors.

To start with, confusion of sense relations, learners may tend to use general terms (superonyms) for more specific terms (hyponyms) or vice versa. To illustrate, the word *economic* (superonym) was intended rather than *financial (hyponym)* in the sentence: "I have economic problems". Also, learners may select the less apt of two related words, or choose the wrong word from a set of near-synonyms. To exemplify, for Turkish EFL learners, the words *nervous* and *excited* are confusing, as the Turkish term *heyecanlı* partially overlaps with both meanings in English. Another category under semantic errors is collocational errors. Collocations refer to habitual associations between words that tend to co-occur in predictable ways. Such associations may be positional, in that one word precedes or follows another (e.g., *heavy rain* rather than **strong rain*); frequency-based, in that certain word combinations are preferred over semantically plausible alternatives (e.g., *make a decision* rather than **do a decision*); or arbitrary, in that the combination cannot be predicted from the meanings of the individual words (e.g., *commit a crime* rather than **perform a crime*). Both confusion of sense relations and collocational errors are included, as the focus is on lexical errors arising from Turkish L1 transfer.

Learners do not benefit solely from recognising errors. Receiving feedback to maintain accuracy in writing is common practice in language classrooms. However, it gives rise to controversial issues about feedback style and sources. It has been argued that teachers in L2 classes may provide inadequate or inaccurate feedback. In addition, external factors, including students' inattention to the feedback given, call into question the effectiveness of error correction (Truscott, 1996, as cited in Hyland & Hyland, 2006).

3. Methodology

3.1. Research Design

This study adopts a case study design, employing a descriptive quantitative method to investigate how teacher and AI-generated written corrective feedback differ in noticing interlingual lexical errors in EFL university writing. The case focuses on a state university English preparatory program and a collection of paragraph and essay writing. Within this case, feedback coverage and type were examined through descriptive quantitative analysis, in which each error instance was compared across three sources of identification:

the learners' retrospective flagging, the teacher's written feedback, and the AI tool's feedback.

3.2. Participants

The participants of this study were EFL students enrolled in the English Preparatory Program at a state university in Türkiye during the 2025-2026 Fall term. All participants shared Turkish as their first language and were placed at the B1 proficiency level at the beginning of the term based on an institutional placement test. The average age of the participants was 18. The study involved two separate groups consisting of different individuals. The opinion paragraph group had six students producing one opinion paragraph, and the opinion essay group included six students generating two opinion essays. Totally 18 texts were formed. Participants were selected through purposive sampling, as only students who attended regularly and completed all required writing tasks were included in the final analysis. Participation was voluntary, and all participants provided written informed consent prior to data collection. Ethical approval was obtained from the Hatay Mustafa Kemal University Ethics Committee, and all participants were assigned pseudonyms to protect their confidentiality.

3.3. Data Collection

Data were collected from the written texts produced by participants during the 2025-2026 Fall term as part of their regular writing instruction. All participants received both teacher-written corrective feedback and AI-assisted feedback as part of their instruction during the term. For this study, groups were arranged around texts that received teacher-written corrective feedback, as teacher feedback constitutes the naturally occurring classroom data. In other words, texts shaped by AI-feedback were excluded from the data. The students in the opinion paragraph group produced one opinion paragraph each, and the students in the opinion essay group produced two opinion essays each, yielding a total corpus of 18 texts. Teacher feedback was provided directly on physical copies of the student texts, which were subsequently digitized by the researcher at the end of the term. To enable comparison, the same texts were analysed by using AI-generated feedback. Specifically, the researcher ran each text through ChatGPT using the identical prompt that students had used during the term. This decision was made deliberately to replicate the original instruction as closely as possible, thereby ensuring consistency and reliability in the comparison between the two feedback sources. It should be noted that the AI feedback generated in this way was produced by the researcher post-hoc and was not shown to students.

As an additional data source, participants were asked to review their Fall term writing texts at the end of the Spring term and highlight the words and sentences they believed had been written under the influence of their first language, Turkish. By the time of this task, participants had been distributed to different classes and had progressed in their English proficiency. This retrospective self-flagging procedure was designed to identify instances of interlingual lexical transfer from the learners' own metalinguistic

perspective, providing a principled basis for locating L1-influenced errors within the corpus. The flagged instances were subsequently classified according to James's (1998) taxonomy of lexical errors, which distinguishes between formal and semantic errors of lexis.

3.4. Procedure

The study was carried out over two terms of the 2025-2026 academic year. During the fall term, which lasted 15 weeks, participants produced their writing texts as part of their regular classroom instruction. The opinion paragraph group wrote one opinion paragraph and received teacher-written corrective feedback on each text. The opinion essay group wrote two opinion essays and similarly received teacher-written corrective feedback. Throughout the term, all participants had access to AI-assisted feedback through ChatGPT, using a standard prompt provided by the researcher. It must be noted that the writing samples were selected from the first drafts, and students were not allowed to use AI in generating the first draft. At the end of the fall term, the physical copies of the student texts were digitized by the researcher.

In the spring term, participants were distributed to different classes as part of the regular program structure. At the end of the spring term, participants were contacted by the researcher and asked to review their fall term texts and retrospectively highlight the words and sentences they believed had been written under the influence of their first language, Turkish. Following data collection, the researcher ran all digitized texts through ChatGPT post-hoc using the identical prompt students had used during the term to generate AI feedback for comparative analysis.

Table 1: The phases, timing, activities, and data collected throughout the study

Phase	Timing	Activity	Data Collected
Phase 1	Fall term (15 weeks)	Opinion paragraph group writes 1 paragraph; receives teacher WCF. Opinion essay group writes 2 essays; receives teacher WCF.	18 student texts with teacher-written corrective feedback
Phase 2	End of Fall term/term break	Researcher digitizes student texts.	Digitized texts;
Phase 3	End of Spring term	Participants retrospectively highlight L1 Turkish transfer instances in Fall term texts.	Learner self-flagged instances
Phase 4	Post-hoc (researcher)	Researcher runs digitized texts through ChatGPT using the original prompt.	AI-generated feedback on all 18 texts

3.5. Data Analysis

Learners highlighted the instances of interlingual lexical transfer on the digitized texts, and errors were classified according to James's (1998) taxonomy of lexical errors. Each instance was then cross-referenced against the teacher's written corrective feedback on the physical paper copies to determine whether the teacher had noticed and addressed it. The same texts were subsequently run through ChatGPT by the researcher using the standard prompt all students had used during the term. AI-generated feedback was

examined in the same manner. Additionally, both teacher and AI feedback were examined for any instances of interlingual lexical errors noticed beyond what learners had flagged, which were recorded as separate instances in the coding table. The type of feedback provided by each source was classified in two ways based on Ellis's (2009) typology of written corrective feedback. The first was whether the correct form was given (direct vs. indirect feedback), and the second was whether the feedback included metalinguistic support, in the form of an error code or an explanation of the error. Each feedback instance was coded on both dimensions, and their combinations formed the feedback categories reported in the results. Feedback classifications were verified against the original teacher-marked texts and ChatGPT outputs.

To organize and manage the data systematically, a coding table was developed in which each error instance constituted a single row. The table recorded the text type, student, task, the flagged instance, its assumed Turkish logic, the James (1998) category, whether each source, learner, teacher, and AI noticed the instance, the type of feedback provided by each source, and the degree of convergence across sources. This structure allowed for a transparent and systematic comparison of the three feedback perspectives across the full corpus of 18 texts. One instance was excluded from the final dataset because the two feedback sources classified it under different error domains (grammatical vs. lexical), preventing unambiguous assignment within the lexical taxonomy; the final corpus comprised 81 error instances. Claude AI was prompted to categorize lexical errors using the study's taxonomy and was consulted for ambiguous cases. The researcher submitted ambiguous instances to Claude AI for independent classification, then reviewed the suggestions and made final coding decisions

4. Results

4.1. Distribution of Interlingual Lexical Error Types

A total of 81 interlingual lexical error instances were identified across the 18 texts produced by the 12 participants: 23 instances occurred in the opinion paragraphs (six texts written by six students) and 58 in the opinion essays (twelve texts written by six students). Figure 1 presents the distribution of these instances according to James's (1998) taxonomy of lexical errors. Semantic errors constituted the overwhelming majority of the dataset (n = 65, 80.2%), with confusion of sense relations (n = 35, 43.2%) and collocational errors (n = 30, 37.0%) emerging as the two most frequent categories. Formal errors were considerably less common (n = 16, 19.8%), comprising misformation (n = 11, 13.6%), and misselection (n = 5, 6.2%). This pattern was consistent across both text types, which suggests that L1 Turkish transfer at the lexical level operated primarily through semantic rather than formal mechanisms in this learner group. In other words, the participants tended to map Turkish semantic and collocational patterns onto English word choices rather than to produce formally deviant word forms.

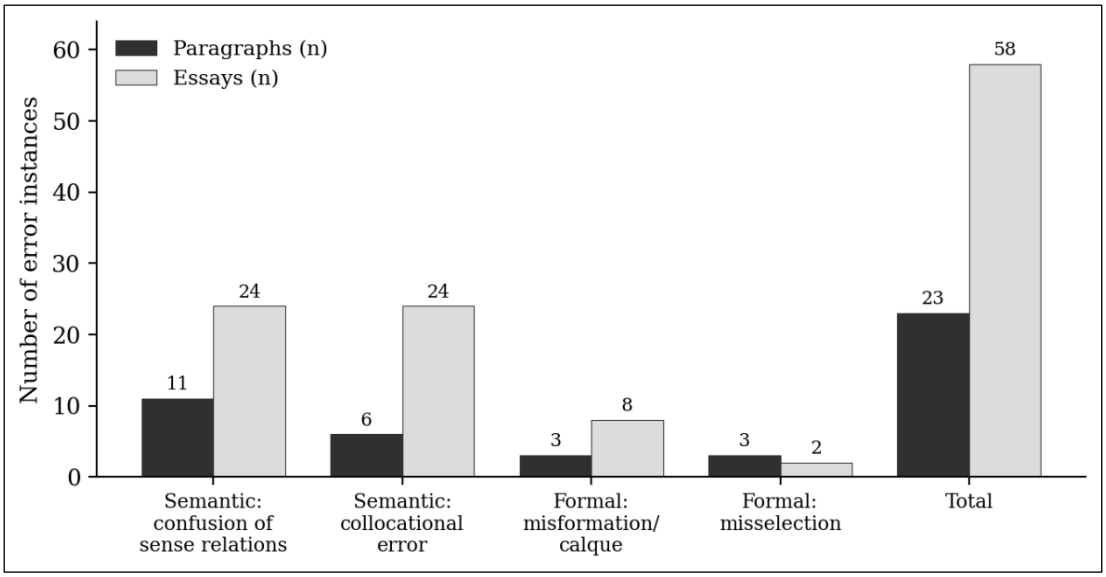

Figure 1: Distribution of Interlingual Lexical Error Types by Text Type (James, 1998)

4.2. Identification of Error Instances by Source

Figure 2 summarizes the 81 error instances identified by the learners themselves, the classroom teacher, and the AI tool. Accordingly, AI feedback addressed by far the largest proportion of instances (n = 69, 85.2%), followed by learner self-flagging (n = 37, 45.7%) and teacher feedback (n = 23, 28.4%). The percentages were calculated over the instances in each text type, and the three sources are not mutually exclusive, as a single instance could be identified by more than one source. This ordering held for both text types, although the teacher's noticing rate was clearly higher in the shorter opinion paragraphs (39.1%) than in the longer opinion essays (24.1%), whereas the AI tool's coverage remained relatively stable across text types (87.0% and 84.5%, respectively).

When the analysis was restricted to the 37 instances that learners had flagged as suspected L1 transfer, the AI tool confirmed 29 of these instances (78.4%), whereas the teacher addressed only 11 (29.7%). This finding suggests that, in most cases, learner intuitions about transfer-induced lexical problems were corroborated by AI feedback but remained unaddressed in teacher feedback.

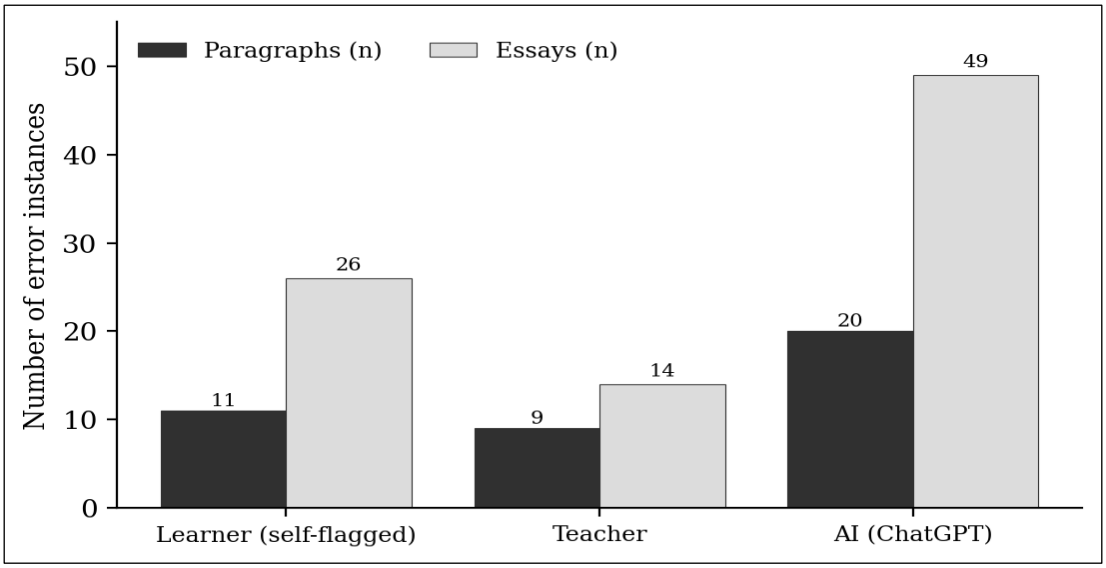

Figure 2: Identification of Interlingual Lexical Error Instances by Source and Text Type

4.3. Convergence of Error Identification Across Sources

Figure 3 presents the convergence patterns among the three sources, derived from the individual noticing records of each error instance. In other words, it shows how many instances were identified by more than one source and how many by a single source only. Full convergence, in which the learner, the teacher, and the AI tool all identified the same instance, was observed in only 11 instances (13.6%). The single largest category consisted of instances identified exclusively by the AI tool (n = 32, 39.5%), which indicates that a substantial share of transfer-induced lexical problems in the dataset would have remained invisible to both the learners and the teacher without AI feedback. Learner–AI convergence without teacher involvement was the second most frequent pattern (n = 18, 22.2%). Notably, no instance was identified by the learner and the teacher while being missed by the AI tool, and only eight instances (9.9%) escaped AI attention while being flagged by the learner alone. Nevertheless, the presence of learner-only (n = 8) and teacher-only (n = 4) instances demonstrates that AI coverage, although extensive, was not exhaustive; certain subtle near-synonym confusions identified by the learner or the teacher were not addressed by the AI tool.

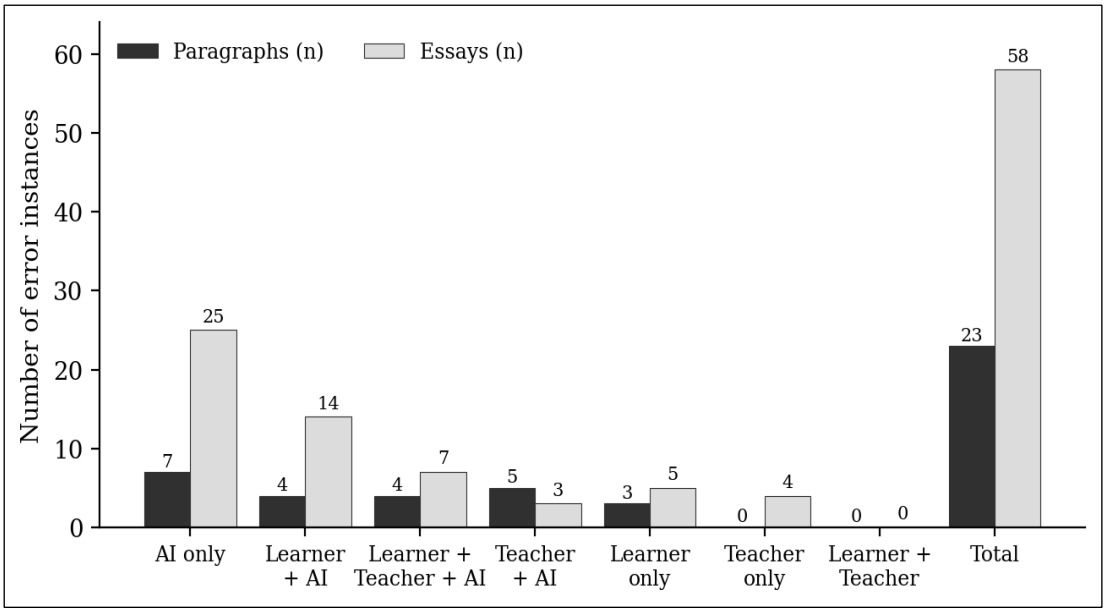

Figure 3: Convergence of Error Identification Across Learner, Teacher, and AI (N = 81)

4.4. Noticing Rates by Error Category

Figure 4 shows the number of errors in each category that were noticed by the learners, the teacher, and the AI tool. Since the categories differ in size, the noticing rates reported below were calculated over the number of instances in each category rather than read from the bar heights. Among the three sources, the AI tool exhibited consistently high identification rates across all four categories, ranging from 80.0% for confusion of sense relations to 100% for formal misselection. The teacher, in contrast, showed a tendency to notice formal errors at higher rates (54.5% for misformation; 60.0% for misselection) than semantic errors (collocational errors 20.0%; sense-relation confusions 22.9%). Learner self-flagging showed a partly similar tendency, with misformation instances flagged most frequently (72.7%).

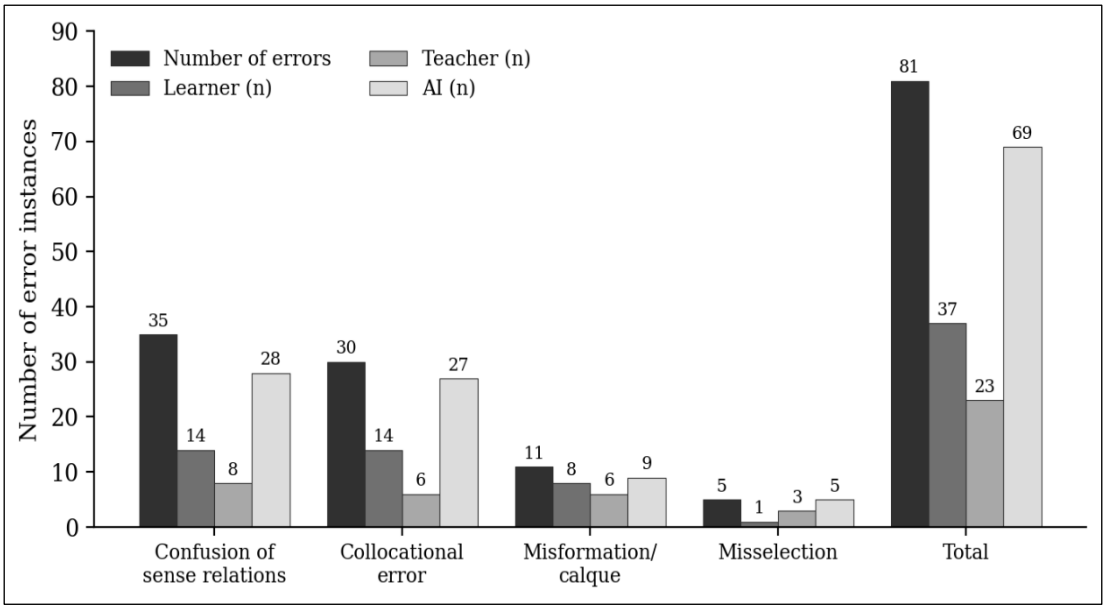

Figure 4: Noticing Rates by Error Category and Feedback Source

4.5. Written Corrective Feedback Strategies

Figure 5 presents the feedback types employed by the teacher and the AI tool, with percentages calculated over the instances each source identified (teacher n = 23; AI n = 69). The feedback was classified according to Ellis's (2009) typology in two dimensions: whether the feedback was direct or indirect, and whether it included metalinguistic support in the form of an error code or an explanation. It was found that the teacher mostly gave indirect feedback with error codes (43.5%), and less frequently indirect feedback with an explanation (21.7%) or direct feedback (26.1% in total). Conversely, the AI tool almost always provided direct feedback with an explanation (98.6%), in other words, direct and metalinguistic feedback combined. The two sources displayed sharply contrasting profiles on the first dimension but a shared reliance on the second. The teacher remained predominantly indirect (17 of 23 instances, 73.9%), most frequently combining indirect indication with an error code (n = 10, 43.5%) or with a brief explanation (n = 5, 21.7%), and supplied the target form in roughly a quarter of the cases (n = 6, 26.1%). Conversely, the AI tool gave direct feedback with an explanation in almost all instances (68 of 69, 98.6%). It usually provided the correct collocation or a suitable near-synonym and explained the usual English usage. In only one instance, there was a correction supplied with a purely evaluative label and no explanation.

A further text-type contrast emerged within the teacher's feedback: in the opinion paragraphs her indirect moves were predominantly explanation-based (5 of 9 instances) and error codes appeared only once, whereas in the essays she shifted almost entirely to code-based marking (9 of 14) with only two explanations, suggesting an economization of feedback effort as text length increased. Finally, the coding notes reveal a distinctive characteristic of the AI feedback. Several lexical errors, particularly collocational ones, were categorized by the tool as grammatical issues or embedded within broader

grammatical commentary, and as a result, the learner may not understand that the problem is lexical rather than grammatical.

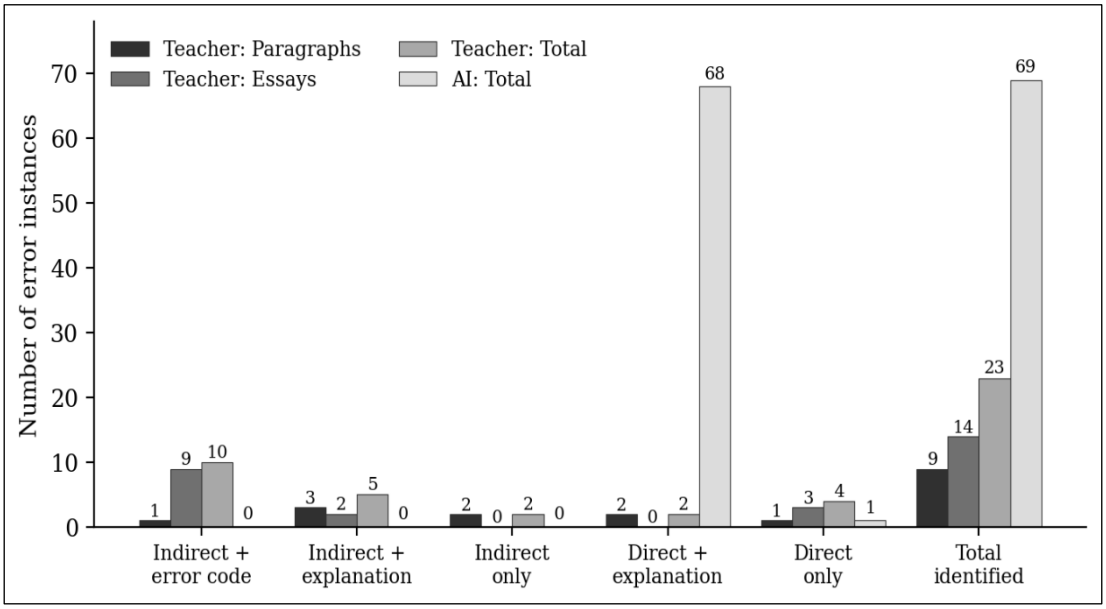

Figure 5: Written Corrective Feedback Strategies as a Share of the Instances Identified by Each Source

4.6. Recurring L1 Transfer Patterns

The recurring errors across the learner texts revealed that different individuals experienced similar interlingual transfer, particularly in verb choices and noun phrases. Figure 6 shows the common expressions mistranslated from the mother tongue (L1 Turkish) and their frequencies. The attempts to translate the verb *edinmek* resulted in four different erroneous forms (e.g., *learn new hobbies*, *take hobby*). Similarly, the verbs *arttırmak* and *geliştirmek* were over-generalized to four inappropriate near-synonyms (e.g., *upgrade your motivation*, *enhance ourselves*). Calques at the noun-phrase level were also systematic. Three different students translated *günlük işler* as *daily work* or *daily routine*, and the false equivalence between *maddi* and *economic* appeared in three instances by three students. Since these patterns were observed in more than one learner, they seem to reflect shared interlanguage processes rather than individual slips. These patterns were not noticed in the same way. All three instances of the *günlük işler* calque were noticed only by the AI tool, and neither the learners nor the teacher identified them. On the other hand, both instances of *reach information* (*bilgiye erişmek*) were flagged by the learners themselves and then confirmed by AI feedback. The *maddi* pattern was in between: the learners flagged two of its three instances. Therefore, it can be said that learner awareness of transfer is selective at the pattern level. Some stabilized calques were not noticed by the learners at all, while others were open to conscious reflection

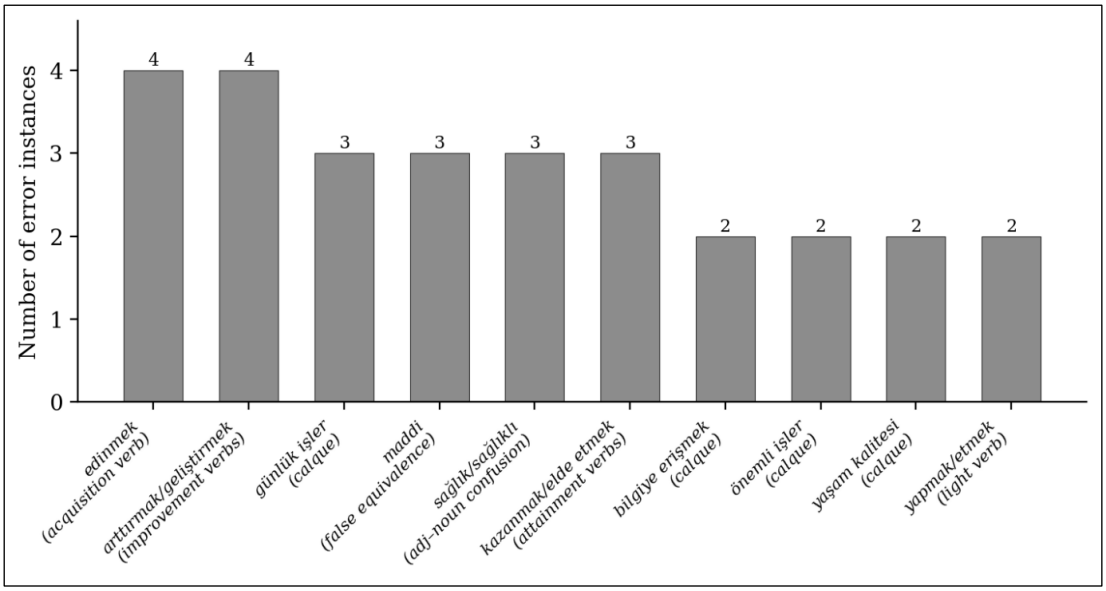

Figure 6: Recurring L1 Turkish Transfer Patterns Across the Dataset

Since the example expressions of each pattern are given in the text above, Figure 6 shows only their frequencies. One error instance may belong to more than one pattern (e.g., *increase your quality of lifestyle* involves both the verb choice and the *yaşam kalitesi* calque); therefore, the frequencies should not be summed.

5. Discussion

The study aimed to explore the noticing rates of interlingual lexical errors by the learners, the classroom teacher, and the AI tool, and to what extent the corrective feedback from the teacher and the AI tool addressed these errors. Several studies reported that interlingual lexical errors including Word for Word translation and collocational errors are among the most frequently encountered error types in learners' writings (Carrió-Pastor & Mestre-Mestre, 2014; Dodigovic *et al.*, 2017; Kırkgöz, 2010; Rattanadilok Na Phuket & Othman, 2015). The results showed that semantic errors were the most frequent interlingual lexical error type at B1 level. This finding is in line with Kazazoğlu (2020), who found that B1-level EFL students made collocation errors that could be attributed to L1 transfer. Similar results were reported in other EFL contexts; for example, translated words from the mother tongue were the most frequent error type among Thai university students (Rattanadilok Na Phuket & Othman, 2015).

Although the proficiency level ranges from A1 to B2 in these studies, dependence on L1 remains stable. For instance, in our case learners started the term at B1 level with opinion paragraph writing. In total, 23 lexical errors were recorded at this level. At the time when they started to write opinion essays they were at early B2 level. However, 58 lexical error instances were identified on essay tasks and the collocational errors and confusion of sense relations remained persistent. Nesselhauf (2005) claimed that collocational errors are dominant in EFL writing regardless of the proficiency level.

Another critical aspect of the study was about noticing errors. The teacher had a lower noticing rate in longer texts. It may be explained by time constraints in a crowded learning setting, which may have reduced the depth of teacher attention to lexical issues. However, the teacher's noticing was not low in every category. The teacher's noticing rate was higher for formal errors than for semantic ones.

The AI tool was the most effective source in identifying errors, outperforming both the learners and the teacher. However, it missed twelve instances (14.8%): eight flagged by the learner alone and four by the teacher alone. Also, its rate was relatively lower for confusion of sense relations. In this sense, it can be said that formal errors such as misformation are more noticeable for human readers, whereas semantic errors, particularly collocational ones, are less visible. Considering that the number of missed instances is small, it can be said that ChatGPT offers reliable writing assistance for teachers and learners. Therefore, as Cengiz and Yigitoglu Aptoula (2026) suggested, AI feedback may complement teacher feedback.

A further noteworthy aspect of the study concerns the learners' retrospective self-flagging, which has received little attention in previous research on interlingual lexical errors and thus constitutes the original contribution of the study. The learners identified nearly half of the error instances (45.7%), a rate higher than that of the teacher, which suggests that learners possess a certain degree of awareness regarding the influence of their L1 on their lexical choices. This finding is consistent with previous research showing that learners are able to identify errors in their own texts (Ferris & Roberts, 2001), even in the absence of external feedback (ElEbyary *et al.*, 2024). The flagged instances matched genuine interlingual lexical errors, which suggests that learner self-flagging, although partial in coverage, seems accurate in what it identifies. The convergence analysis supports this interpretation. All three sources identified the same instance in only 11 cases (13.6%), and the second most frequent pattern was the instances identified by the learner and the AI tool but not by the teacher (n = 18, 22.2%). It seems that each source notices different types of errors, and no source identifies all the errors on its own. Therefore, it can be said that error identification in EFL writing works best as a complementary process, in which learner awareness, teacher judgement, and AI assistance each cover a different part of the errors.

As for feedback delivery, the two sources differed not in whether they gave metalinguistic support but in what the support was used for. The teacher mostly gave indirect feedback with error codes, which withholds the correct form and directs the learner to self-correction. The AI tool, conversely, gave the correct form together with an explanation in almost all instances. This maximizes explicitness, but it may reduce the learner's engagement in correcting the error. In addition, the AI tool sometimes presented lexical errors as grammar issues, and in one instance its correction changed the student's intended meaning. In this sense, the explicitness of AI feedback does not guarantee that the learner understands the lexical nature of the problem.

The study has certain limitations. It was conducted with a small group of participants (n = 12) in a single institutional context, with one teacher as the human

feedback source; therefore, the findings cannot be generalized to other settings. In addition, the self-flagging procedure was retrospective in that the learners reviewed their texts at the end of the following term, by which time their proficiency had increased. In this sense, the noticing rates reflect the learners' awareness at the time of review rather than at the time of writing. Moreover, the AI outputs were obtained at a single point in time, and ChatGPT may generate different responses to the same texts. Despite these limitations, the findings offer practical implications: AI tools may be employed to extend error coverage, particularly for semantic errors that remain less visible to human readers, while teacher feedback retains its value in addressing subtle, context-dependent problems and in engaging learners in self-correction.

Acknowledgement
AI tools (Claude, Anthropic) were used for language editing and drafting support; all analyses, findings, and interpretations are the author's own.

Conflict of Interest Statement
The author declares no conflict of interest regarding the research, authorship, or publication of this article.

About the Author
Funda Dündar is an EFL instructor at Eskisehir Technical University, School of Foreign Languages and a PhD candidate at Hatay Mustafa Kemal University, Türkiye, in Informatics Department. Her research interests include English Language Teaching and EFL writing.
Academic profile: https://akademik.eskisehir.edu.tr/fundadundar
Google Scholar: https://scholar.google.com/citations?user=GdT40_UAAAAJ

References

Al-Olimat, S. I. (2015). Using computer-mediated corrective feedback modes in developing students' writing performance. *Teaching English with Technology, 15*(3), 3–30. Retrieved from https://files.eric.ed.gov/fulltext/EJ1138425.pdf

Arts, J. G., Jaspers, M., & Brinke, D. J. (2016). A case study on written comments as a form of feedback in teacher education: So much to gain. *European Journal of Teacher Education, 39*(2), 159–173. https://doi.org/10.1080/02619768.2015.1116513

Bhela, B. (1999). Native language interference in learning a second language: Exploratory case studies of native language interference with target language usage. *International Education Journal, 1*(1), 22–31.

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. Retrieved from https://files.eric.ed.gov/fulltext/EJ1042860.pdf

Cengiz, Y., & Yigitoglu Aptoula, N. (2026). Comparing AI and human feedback at higher education: Level appropriateness, quality and coverage. *Language Teaching, 59*(1), 122–127. https://doi.org/10.1017/S0261444825000199

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching, 5*(4), 161–170. https://doi.org/10.1515/iral.1967.5.1-4.161

Council of Europe. (n.d.). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Structured overview of all CEFR scales*. https://rm.coe.int/168045b15e

Crosthwaite, P., & Sun, S. (2025). Generative AI and L2 written feedback studies: A scoping review. *RELC Journal*. https://doi.org/10.1177/00336882251386530

Dodigovic, M., Ma, C., & Jing, S. (2017). Lexical transfer in the writing of Chinese learners of English. *TESOL International Journal, 12*(1), 75–90. Retrieved from https://files.eric.ed.gov/fulltext/EJ1247836.pdf

ElEbyary, K., Shabara, R., & Boraie, D. (2024). The differential role of AI-operated WCF in L2 students' noticing of errors and its impact on writing scores. *Language Testing in Asia, 14*(1). https://doi.org/10.1186/s40468-024-00312-1

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal, 63*(2), 97–107. https://doi.org/10.1093/elt/ccn023

Er, H. K., & Küçükali, E. (2024). The Effects of Face-to-Face vs. Digital Feedback in an EFL Writing Context: Comparison of Two Turkish State Universities. Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 24(1), 389-411. https://doi.org/10.17240/aibuefd.2024..-1340007

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing, 10*(3), 161–184. https://doi.org/10.1016/S1060-3743(01)00039-X

Ferris, D., Liu, H., & Rabie, B. (2011). "The job of teaching writing": Teacher views of responding to student writing. *Writing & Pedagogy, 3*(1), 39–77. Retrieved from https://doi.org/10.1558/wap.v3i1.39

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course* (2nd ed.). Lawrence Erlbaum Associates. Retrieved from https://bpb-us-e2.wpmucdn.com/websites.umass.edu/dist/c/2494/files/2015/08/Gass.Second-Language-Acquisition.pdf

Gök, Ş. (2020). Contrastive error analysis of Turkish EFL learners in writing. *International Journal of Language and Literary Studies, 2*(4), 236–242. https://doi.org/10.36892/ijlls.v2i4.429

Guo, K., & Wang, D. (2024). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies, 29*, 8435–8463. https://doi.org/10.1007/s10639-023-12146-0

Han, T., & Sari, E. (2024). An investigation on the use of automated feedback in Turkish EFL students' writing classes. *Computer Assisted Language Learning, 37*(4), 961–985. https://doi.org/10.1080/09588221.2022.2067179

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing, 30*, 31–44. https://doi.org/10.1016/j.jslw.2015.08.002

Hasırcıoğlu, S., & Öztürk, M. S. (2024). Analyzing mother tongue interference in Turkish EFL students' written products: A lexical and syntactical perspective. *Öğretim Teknolojisi ve Hayat Boyu Öğrenme Dergisi*. https://doi.org/10.52911/itall.1438891

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83–101. https://doi.org/10.1017/S0261444806003399

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. Longman. Retrieved from https://books.google.ro/books/about/Errors_in_Language_Learning_and_Use.html?id=DLBoAAAAIAAJ&redir_esc=y

Kaliisa, R., Misiejuk, K., López-Pernas, S., & Saqr, M. (2026). How does artificial intelligence compare to human feedback? A meta-analysis of performance, feedback perception, and learning dispositions. *Educational Psychology, 46*(1), 80–111. https://doi.org/10.1080/01443410.2025.2553639

Kazazoğlu, S. (2020). The impact of L1 interference on foreign language writing: A contrastive error analysis. *Journal of Language and Linguistic Studies, 16*(3), 1177–1188. https://doi.org/10.17263/jlls.803621

Kırkgöz, Y. (2010). An analysis of written errors of Turkish adult learners of English. *Procedia – Social and Behavioral Sciences, 2*(2), 4352–4358. https://doi.org/10.1016/j.sbspro.2010.03.692

Liu, Y. (2023). A comparison of automated corrective feedback and traditional corrective feedback: A review study. *The Educational Review, USA, 7*(9), 1365–1368. https://doi.org/10.26855/er.2023.09.024

Makino, T. (1993). Learner self-correction in EFL written compositions. *ELT Journal, 47*(4), 337–341. Retrieved from https://eric.ed.gov/?id=EJ474539

Mao, S., & Crosthwaite, P. (2019). Investigating written corrective feedback: (Mis)alignment of teachers' beliefs and practice. *Journal of Second Language Writing, 45*, 1–14. https://doi.org/10.1016/j.jslw.2019.05.004

Nesselhauf, N. (2005). *Collocations in a learner corpus*. John Benjamins. Retrieved from https://books.google.ro/books/about/Collocations_in_a_Learner_Corpus.html?id=RasDwVvBp00C&redir_esc=y

Pham, L. N. (2023). *The interplay between learner autonomy and indirect written corrective feedback in EFL writing*. *TESOL Journal*, 14. https://doi.org/10.1002/tesj.694

Rattanadilok Na Phuket, P., & Othman, N. B. (2015). Understanding EFL students' errors in writing. *Journal of Education and Practice, 6*(32), 99–106. Retrieved from https://files.eric.ed.gov/fulltext/EJ1083531.pdf

Richards, J. C. (1971). A non-contrastive approach to error analysis. *English Language Teaching Journal, 25*(3), 204–219. [if kept, change the in-text year from 1972 to 1971]

Son, M. (2024). L2 language development in oral and written modalities. *Studies in Second Language Acquisition*, *46*(3), 841–868. https://doi.org/10.1017/S0272263124000329

Wang, D. (2024). Teacher- versus AI-generated (Poe application) corrective feedback and language learners' writing anxiety, complexity, fluency, and accuracy. *International Review of Research in Open and Distributed Learning, 25*(3), 37–56. https://doi.org/10.19173/irrodl.v25i3.7646

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing, 21*(4), 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Zheng, Y., Yu, S., Liu, X., & Gu, M. (2018). Student engagement with teacher-written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing, 36*, 35–47. https://doi.org/10.1016/j.asw.2018.03.001